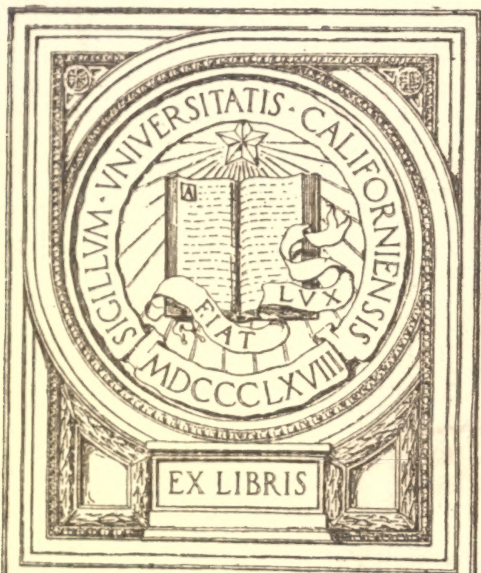


GIFT OF
JULIUS WANGENHEIM '87



EX LIBRIS

Mathematics Dept.

H. Hall

Magd: Coll Camb.

A

TREATISE

ON

PROBABILITY:

FORMING

THE ARTICLE UNDER THAT HEAD IN THE SEVENTH EDITION OF

THE ENCYCLOPÆDIA BRITANNICA.

BY

THOMAS GALLOWAY, M.A. F.R.S.

SECRETARY OF THE ROYAL ASTRONOMICAL SOCIETY.

EDINBURGH:

ADAM AND CHARLES BLACK, NORTH BRIDGE,

BOOKSELLERS TO HER MAJESTY FOR SCOTLAND.

1839.

6712
G-34
Math
dept.

NO. VIII
ANNUAL

gift to J. Wangerheim
to Math. Dept.

EDINBURGH:
BALFOUR & JACK, Printers, Niddry Street.

CONTENTS.

INTRODUCTION.....	Page 1.
-------------------	---------

SECTION I.

GENERAL PRINCIPLES OF THE THEORY OF PROBABILITY.....	P. 16.
--	--------

ARTICLES.—1-4. Definition of the terms Chance and Probability. 5, 6. Measure of mathematical probability. 7. Probability of compound events. 8. Example in numbers. 9. Probability of an event which may happen in different ways. 10. Identity of the formulæ for simultaneous and successive events.

SECTION II.

OF THE PROBABILITY OF EVENTS DEPENDING ON A REPETITION OF TRIALS, OR COMPOUNDED OF ANY NUMBER OF SIMPLE EVENTS, THE CHANCES IN RESPECT OF WHICH ARE KNOWN <i>A PRIORI</i> AND CONSTANT.....	P. 27.
---	--------

ARTICLES.—11, 12. Probabilities of the different combinations that may happen in a series of trials. 13. Application of the binomial theorem. 14. Example in a particular case. 15. Extension to the case of any number of simple events. 16—21. Examples. 22, 23. Artifice for abbreviating the calculation. 24, 25. Determination of

the number of trials required to render the probability of an event equal to an assigned fraction.

SECTION III.

OF THE PROBABILITY OF EVENTS DEPENDING ON A REPETITION OF TRIALS, OR COMPOUNDED OF ANY NUMBER OF SIMPLE EVENTS, THE CHANCES IN RESPECT OF WHICH ARE KNOWN *A PRIORI*, AND VARY IN THE DIFFERENT TRIALS.....P. 51.

ARTICLES.—26. Expression for the probability of the different possible results in particular cases. 27, 28. Extension to the general case. 29. Examples. 30. Solution of a question proposed by Huygens.

SECTION IV.

OF MATHEMATICAL AND MORAL EXPECTATION.....P. 59.

ARTICLES.—31. Definition of the term expectation. 32. Example of mathematical expectation. 33. Mathematical expectation not applicable in particular cases. 34. Hypothesis of Bernoulli. 35, 36. Formulæ to express the value of a moral expectation. 37—39. Consequences of the formulæ. 40. Application of the theory of Bernoulli to the subject of insurances. 41. Petersburg problem.

SECTION V.

OF THE PROBABILITY OF FUTURE EVENTS DEDUCED FROM EXPERIENCE.....P. 75.

ARTICLES.—42. Hypotheses respecting the causes of an event. 43. Determination of the probabilities of the different hypotheses. 44. Probability of a future event deduced from the probabilities of the hypotheses. 45. Extension of the formulæ to any finite number of causes or hypotheses. 46. Sense in which the term *cause* is used in

this theory. 47, 48. Examples of the formulæ. 49. Case in which the different causes are not equally probable. 50. Application of the formulæ to physical or moral events. 51, 52. Extension of the formulæ to the case of an infinite number of different causes.

SECTION VI.

OF BENEFITS DEPENDING ON THE PROBABLE DURATION OF HUMAN LIFE..... P. 90.

ARTICLES.—53. Principles on which the probability of life is computed. 54, 55. Method of computing the value of an annuity on a single life. 56. Of a life annuity for terms of years. 57. Of annuities on joint lives. 58. On the survivor of any number of lives. 59—63. Methods of computing the values of assurances on lives.

SECTION VII.

OF THE APPLICATION OF THE THEORY OF PROBABILITY TO TESTIMONY, AND TO THE DECISIONS OF JURIES AND TRIBUNALS..... P. 100.

ARTICLES.—64, 65. Expression for the probability of an event attested by a single witness on an assumed hypothesis. 66. Case in which the event attested is extremely improbable. 67. Case in which the character of the witness is altogether unknown. 68. Expression of the probability of an attested event, regard being had to the *a priori* probability of the event. 69, 70. Probability of events attested by several witnesses. 71. Formulæ for the case of conflicting testimony. 72—75. Successive testimony, or tradition. 76. Application to the verdicts of juries. 77—81. Probability of acquittal and condemnation under different hypotheses. 82, 83. Probability of a verdict being correct when pronounced by a given majority. 84. Numerical expression for the error of a verdict when arbitrary values are given to the constants. 85, 86. Values of the constants, deduced from the records of the criminal courts in France.

SECTION VIII.

OF THE SOLUTIONS OF QUESTIONS INVOLVING LARGE NUMBERS. P. 128.

ARTICLES.—87–89. Form of the expression for the probability of the most probable event in a large number of trials. 90. Stirling's theorem for approximating to its numerical value. 91. Example of the formula. 92. Probability of any assigned result of a large series of trials. 93. Example of the rapid diminution of the terms of the binomial towards the beginning and end of the series when the exponent is large. 94. Probability of the number of occurrences of an event falling within given limits. 95. Determination of the limits for an assigned probability. 96. Method of computing the value of the integral which expresses the probability. 97. General consequences of the formulæ. 98. Method of using the table. 99. Application to numerical examples. 100–102. Extension to the case of probabilities deduced from experience. 103. Application to a question of mortality. 104. To the determination of a physical cause. 105. To determine the limits of loss or gain in a multitude of pecuniary risks. 106. General deductions relative to the ultimate results of play. 108, 109. Digression relative to a question respecting the duration of play.

SECTION IX.

OF THE MOST PROBABLE MEAN RESULTS OF NUMEROUS DISCORDANT OBSERVATIONS, AND THE LIMITS OF PROBABLE ERROR.....P. 164.

ARTICLES.—110. Object of the application of the theory. 111. Statement of the problem. 112–123. Solution of the problem, and expression for the probability of the average of a large number of observations falling within assigned limits. 124. Connection of the special constant to which the average continually approaches, with the centre of gravity of a curve line. 125–126. Cases in which the number of possible values of the phenomenon is limited.

127—130. Examples when the law of the probability is supposed to be known. 131—132. Form of the function which expresses the law of probability, on the hypotheses of positive and negative errors being equally probable. 133. Illustration by means of a curve. 134. Determination of the constants on this hypothesis. 135—142. Investigation of an expression for the constant of the law of facility, in terms of the results of observations already made. 143. Different expressions for the limits. 144. Use of the formulæ in determining the relative values of tables. 145—148. Expressions for the precision, weight, probable error, and mean error of the results of observations. 149. Determination of the most probable value of a function of several quantities, separately determined by observation, with the precision and weight of the determination.

SECTION X.

OF THE METHOD OF LEAST SQUARES..... P. 200.

ARTICLES.—151—154. Method of forming equations of condition. 155—157. Determination of the most probable values of the elements from the equations of condition, and the probable amount of the remaining errors. 158. Agreement of this solution with the method of least squares. 159. Example in the case of a single element. 160. Example in the case of several elements.

*The Reader is requested to substitute the following List of
Errata for that at the end of the volume.*

ERRATA.

Page 40, note, for *Essai sur les Probabilités*, read *Théorie Analytique des Probabilités*.

- 43, line 7 from foot, for *adopt* read *adapt*.
- 49, line 13, read beginning *with*. In the note, read *Traité Élémentaire*.
- 54, line 7, for $a-m$, read $a-m'$.
- 81, first line, for $(1+2+3...+n)$, read $(1+2+3...+n)$.
- 89, line 9, for $m-m$, read $m+m'$.
- 135, last line, for $(n+x)^{-n+x-\frac{1}{2}}$, read $(n+x)^{-n-x-\frac{1}{2}}$.
- 188 line 8, for e^{-2} , read $e^{-\theta^2}$.
- 193, line 8, delete *products of the*.
- 198, line 9, for e' , e'' , read $e'e''$.
- 203, line 19, for hc' , read hc .
- — line 3 from foot, for (1), read (2).
- 209, last line, for *les* read *des*.
- 210, line 2 from foot, for 1833, 1834, and 1835, read 1834, 1835, and 1836.
- 211, last line, for *Abhandlungen*, read *Nachrichten*.

P R E F A C E.

IN drawing up the following treatise for the Encyclopædia Britannica, my design has been, to present a general view of the principles, applications, and more important results of the mathematical theory of Probability, as laid down in the best and most recent works on the subject; particularly those of Laplace and Poisson: and, without entering into the details of mathematical difficulties, to explain the methods of applying analysis to the solution of the principal questions, as fully as the limits of the space which could be appropriated to the article would permit.

In the prosecution of this design, questions connected with lotteries and games of chance, the subjects to which the earlier writers on Probability chiefly confined themselves, and which are frequently supposed to form a principal part, if not the whole of the science, occupy but a small portion of the work; indeed they are only introduced as furnishing exam-

ples to illustrate the principles or the formulæ. The theory of moral expectation has been developed at some length, because it admits of important applications in the affairs of life. The subject of mortality, and the computation of annuities and other pecuniary interests depending on life contingencies, (the most common application of the theory,) having been treated in detail under other heads in the *Encyclopædia*, scarcely came within the scope of the article ; but the method of applying the principles of the science to the data furnished by the mortality tables, has been explained, and the formulæ of most frequent practical use have been deduced. The section on the probability of testimony and the verdicts of juries, is more expanded, and contains formulæ which resolve the principal questions the subject presents, having regard to the different causes of uncertainty. In the following section, formulæ are given for obtaining approximate values of expressions which involve very large numbers. Such expressions occur in all the higher and more interesting applications of the doctrine of chances, and give rise to the greater part of the mathematical difficulties which attend the subject. The probability of the result of a large number of experiments or observations being contained within given limits, is expressed by a definite integral, which occurs in various other applications of analysis ; in order to facilitate the computation of such probabilities, a table of the values of this inte-

gral is given at the end, by which, and by a simple arithmetical operation, we are enabled to solve a multitude of questions of the most interesting and otherwise difficult kind. In the investigation of the most probable mean value of a quantity, to be determined in magnitude or position, from a series of observations liable to error, and the determination of the limits of probable uncertainty, I have followed the very general and elegant analysis of Poisson. The last section contains the application of the general theory to the determination of the most advantageous method of combining equations of condition, and an explanation of the celebrated rule of least squares, which is shewn to coincide with that pointed out by the theory as giving the most probable values of the elements involved.

The principal application which has hitherto been made of the theory explained in the last two sections, has been to questions of astronomy; but it may be applied, with equal advantage, to every branch of physical inquiry, or statistical research, where the object is to deduce precise and accurate mean results from multiplied observations, or to detect the existence of general laws, or the operation of constant causes, among large masses of phenomena.

Although considerable pains have been taken to treat the subject in an elementary manner, it was impossible, in some places, particularly the three last sections, to avoid discussions which belong to the

more abstruse parts of the mathematics. The higher and more useful parts of the theory, do not admit of being treated in any other way, unless the reader is prepared to abandon the accuracy of exact science, and to accept of assertion for proof. I have endeavoured, however, to place the demonstrations in their clearest light, and to give the expressions in their simplest forms; and it is hoped there is nothing in the present work beyond the reach of those who have a moderate knowledge of the differential and integral calculus, and are acquainted with the first principles of the theory of finite differences.

It may be proper, perhaps, to give a reason for deviating from the usual notation, by writing fractional expressions in the form of ratios or quotients. This practice was adopted for the purpose of economising space in printing, when it was perceived that the article was likely to exceed the prescribed limits. It may be considered objectionable on the ground of its being less familiar to the reader; but as it is here used it can hardly occasion any real embarrassment.

LONDON,
February 12, 1839.

PROBABILITY.

THE doctrine of probability is an extensive and very important branch of mathematical science, the object of which is to reduce to calculation the reasons which we have for believing or expecting any contingent event, or for assenting to any conclusion which is not necessarily true. When it is considered that the whole edifice of human science, with the exception of a few self-evident truths, such as the axioms of geometry, is nothing more than an assemblage of propositions which can only be pronounced to be more or less probable, the importance of a calculus which enables us to appreciate exactly the degree of probability existing in each case, will be readily understood.

Our reasons for judging an event to be probable or improbable, are derived from two distinct sources ; first, an *a priori* knowledge of the causes or circumstances which determine its occurrence ; and, secondly, when the causes are unknown, experience of what has already happened in the same circumstances, or in circumstances apparently similar. Suppose, for example, a hundred white balls to be

placed in an urn along with fifty black balls, and that a person, blindfold, proceeds to draw a ball, there is to us, who are acquainted with the contents of the urn, a determinate probability that the ball which is drawn will be white. The balls being supposed to be all in precisely the same circumstances with respect to facility of drawing, we assume that there is the same chance of drawing any one ball as of drawing any other ; and, consequently, since there are two white balls for each black ball, and therefore two chances of drawing one of the first colour for each chance of drawing one of the second, we conclude the event which consists in the drawing of a white ball to be twice as probable as the opposite event, or the drawing of a black ball. In this case our knowledge of the contents of the urn enables us to judge of the probable result of the drawing. Suppose, however, that antecedently to the drawing, we were entirely ignorant of the contents of the urn, but that after a great number of trials have been made, (the ball drawn being always replaced in the urn after each trial, in order that the circumstances may be the same in all the trials) it has been observed that a white ball has been drawn twice as often as a black ball, we presume that the urn contains twice as many white as black balls, and consequently affords twice as many chances of drawing a white ball as of drawing a black ; and this presumption becomes stronger in proportion to the number of instances included in the observation. In this case experience makes up for the want of *a priori* knowledge, and affords a measure of the probability of the result of a future trial.

It is only in a comparatively small number of cases that all the possible ways in which an event may happen are known *a priori*, and in which, consequently, the ratio of the

number of chances favouring the event to the whole number of existing chances is determinate. In fact, most of the questions of this class to which the calculus can be applied, are connected with lotteries and games of hazard. The results obtained from the analysis of such questions cannot be considered as being of any great value in themselves, but they frequently throw light on subjects of far higher importance which present analogous combinations. It is true that the mathematical theory comes in aid of moral considerations, and demonstrates the ruinous tendency of gambling even when the conditions of the play are equal, mathematically speaking; but, unfortunately, those who indulge a passion for this vice are seldom capable of appreciating the force of such arguments. The principal advantage which has resulted from the application of analysis to games of chance is the extension and improvement of the calculus to which it has led.

The calculation of the probabilities of events, the chances of which are not known *a priori*, but inferred from experience, is founded on the presumed constancy of the laws of nature, in obedience to which events depending on constant though unknown causes, are always reproduced in the same order when considered in large numbers. Among the various phenomena of the physical and moral world, nothing is more remarkable than the constancy which is observed to prevail in the recurrence of events of the same kind. The ratio of male to female births furnishes a noted instance. If we consider only a small number of births, nothing can be more uncertain than the result; but taking a very large number, as those of a whole kingdom in the course of a year, the proportion of males to females is found to be almost inviolable, and nearly as 21 to 20. The mean duration of human life affords another familiar example. Notwithstand-

ing the proverbial uncertainty of life, the differences of constitutions, and the various accidents to which mankind are exposed, the average duration of the lives of a large number of individuals living in the same country is always found to be very nearly the same, insomuch that pecuniary risks depending on it, if undertaken in sufficiently large numbers, are among the least uncertain of all commercial speculations. A similar constancy is remarked in the results of statistical inquiries of every kind. The number of crimes of the same species committed in a year, the ratio of the number of acquittals to the number of trials, the number of conflagrations, of ships lost in a particular trade, of letters which pass through the post-office, of patients admitted into the public hospitals; in every case the numbers in a given time are observed to fluctuate between very narrow limits, and to approach nearer and nearer, as the observation is more extended, to fixed mean values.

This constant approximation to fixed ratios, which is proved by all experience, in the recurrence of events of the same kind, enables us to apply the calculus of probabilities to many of the most interesting questions connected with our social and political institutions; and to determine the average result of a series of coming events with as much precision as if their chances were determinate, and known *a priori*, like that of obtaining a given point with the throw of a die. Whatever be the nature of the phenomenon under consideration, whether it belong to the physical or moral order of things, the calculus is equally applicable when the requisite data have been determined from experience.

The foundations of the mathematical theory of probabilities were laid by Pascal and Fermat about the middle of the 17th century. Among some other questions relating to chances, the following was proposed to Pascal. "Two

persons sit down to play on the condition that the one who first gains three games shall be the winner of the stakes. The first having gained two games, and the second one, they agree to leave off and divide the stakes in proportion to their respective probabilities of winning: what share is each entitled to take?" Pascal solved the question, but by a method which was applicable only to the particular case. Fermat, to whom it was communicated by Pascal, employed the direct and general method of combinations, and gave a solution which could be applied to the case of any number of players. His reasoning, however, did not at first appear to Pascal to be satisfactory, and a correspondence on the subject took place between these two illustrious geometers, which is preserved in their respective works, and throws some light on the history of mathematics in that age.¹

About the same period Huygens composed his tract *De Ratiociniis in Ludis Aleæ*, which was first published in the *Exercitationes Geometricæ* of Schooten in 1658. This was the first systematic treatise which appeared on the doctrine of chances. It contained an analysis of the various questions which had been solved by Pascal and Fermat, and at the end five new questions were proposed, the solutions of which, simple as they may now appear, were then attended with considerable difficulty. The analysis of two of them was in fact given for the first time by Montmort, half a century after their publication. Huygen's tract was translated into English and published in 1692, with some additional remarks relative to the advantage of the banker in the game of Pharaon, in an *Essay on the Laws of Chance*, edited and

¹ *Œuvres de Blaise Pascal*, tome iv. Paris, 1819; *Opera Petri de Fermat*, Tolosæ, 1679.

supposed to have been written by Motte, then Secretary of the Royal Society.

James Bernoulli appears to have been the first who perceived that the theory of probability may be applied to much more important purposes than to regulate the stakes and expectations of gamblers, and that the phenomena, both of the moral and physical world, anomalous and irregular as they appear when viewed in detail, exhibit, when considered in large numbers, a constancy of succession which renders their occurrence capable of being submitted to numerical estimation. The *Ars Conjectandi*, published in 1713, seven years after the death of the author, contains a number of interesting questions relative to combinations and infinite series ; but the most remarkable result which it contains is a theorem respecting the indefinite repetition of events, which may be said to form the basis of all the higher applications of the theory. It consists in this, that if a series of trials be instituted respecting an event which must either happen or fail in each trial, the probability becomes greater and greater, as the number of trials is increased, that the ratio of the number of times it happens, to the whole number of trials, will be equal to its *a priori* probability in a single trial ; and that the number of trials may be made so great as to give a probability, approaching as nearly to certainty as we please, that the difference between the ratio of its occurrences to the number of trials, and the fraction which measures its *a priori* probability, will be less than any assigned quantity. Bernoulli informs us, that the solution of this important theorem had engaged his attention during a period of twenty years.

In the interval between the death of Bernoulli and the appearance of the *Ars Conjectandi*, Montmort published

his *Essai d'Analyse sur les Jeux de Hazard*. The first edition was in 1708; the second, which is considerably extended, and enriched by several letters of John and Nicolas Bernoulli, appeared in 1713. The work possesses considerable merit; but being chiefly confined to the examination of the conditions of games of chance, many of which are now forgotten, it has lost much of its original interest.

About the same time, Demoivre began to turn his attention to the subject of probability, and his labours, which were continued during a long life, contributed greatly to the advancement of the general theory, as well as the extension of some of its most interesting applications. Demoivre's first publication on the subject was a Latin memoir *De Mensura Sortis*, in the *Transactions of the Royal Society* for 1711. His *Essay on the Doctrine of Chances* first appeared in 1716; a second edition in 1738; but the third and most valuable, including also his *Treatise on Annuities on Lives*, is dated 1756. This work contains a great variety of questions relating to chances, solved with much clearness and elegance; but it is chiefly remarkable for the theory of recurring series, there given for the first time, which is of important use in investigations of this kind, and is in fact equivalent to the methods employed in the modern calculus for the integration of equations of finite differences having constant co-efficients. Of the particular results obtained by Demoivre, one of the most important in reference to theory, is an extension of the theorem of James Bernoulli, above mentioned. It follows from Bernoulli's theorem, that if we have a given probability that the ratio of the number of occurrences of an event to the whole number of trials, will approach to the *a priori* probability of the event within certain given limits, those limits will become narrower and

narrower, as the number of trials is multiplied ; but in order to complete the theorem, it is necessary to assign the numerical value of the probability that in a large number of future trials, the number of occurrences will fall within assigned limits. For this purpose we must find the product of the natural numbers 1, 2, 3, 4, &c., up to the number of trials ; an operation which, if attempted by direct multiplication, becomes very laborious, even when the number of trials is inconsiderable, and when the number is great, as 10,000 for example, is altogether beyond the reach of human industry. A formula was however discovered by Stirling, by means of which an approximate value of the product is found by the summation of a few of the first terms of a series which converges the more rapidly as the number of trials is greater. With the aid of this formula, Demoivre was enabled to assign the probability in question, and thus give a practical value to the theorem of Bernoulli.

The objects and important applications of the theory of probabilities having been made known by the works now mentioned, the subject has ever since been regarded as one of the most curious and interesting branches of mathematical speculation, and accordingly has received more or less attention from almost every mathematician of eminence. A great variety of questions connected with it and especially relating to lotteries, are interspersed in the volumes of the *Paris* and *Berlin Memoirs*, (particularly the latter,) by John and Nicolas Bernoulli, Euler, Lambert, Beguelin, and others. D'Alembert has likewise treated of the theory in several of the volumes of his *Opuscula* ; and it is not a little remarkable, that in some instances its first principles should have been misunderstood by so ingenious and profound a writer. In the *St. Petersburg Memoirs*, (vol. v.) there is an interesting

paper by Daniel Bernoulli on the relative values of the expectations of individuals who engage in play, or stake sums on contingent benefits, when regard is had to the difference of their fortunes ; a consideration which, in many cases, it is necessary to take into account ; for it is obvious, that the value of a sum of money to an individual, depends not merely on its absolute amount, but also on his previous wealth. On this principle Bernoulli has founded a theory of *moral expectation*, which admits of numerous and important applications to the ordinary affairs of life. The *Transactions of the Royal Society* for the years 1763 and 1764, contain two papers by the Rev. Mr. Bayes, with additions to the latter by Dr. Price, which deserve to be noticed, inasmuch as the principles on which the probability of an event is determined, when the event depends on causes of which the existence and influence are only presumed from experience, are there for the first time correctly laid down. The question proposed and solved by Bayes was this : a series of experiments having been made relative to an event, to determine the presumption there is, that the fraction which measures its probability falls within given limits.

One of the earliest applications of the theory of probability was to determine, from observations of mortality, the average duration of human life, and the value of pecuniary interests depending on its continuance or failure. This particular application appears to have been first thought of, or at least attempted to be carried into practical effect, in Holland, by Hudde and the celebrated pensionary De Witt ; but the first tables of mortality, with the corresponding values of annuities on single lives, were constructed by our illustrious countryman Dr. Halley, and published in the *Philosophical Transactions* for 1693. For the history of

this branch of the subject, we refer to the two articles, ANNUITIES and MORTALITY in the seventh edition of the *Encyclopædia Britannica*. We may remark, however, that although the English writers, who have expressly treated of it, have almost without exception confined themselves to the explanation of the methods of computing annuity tables, and of determining from them the values of sums depending on life contingencies, the aid which this branch of economy derives from the general theory of probabilities, is by no means confined to the consideration of such elementary questions. The number of observations necessary to inspire confidence in the tables, the extent to which risks may be safely undertaken, the comparative weights of different sets of observations, and the probable limits of departure from the average results of previous observations in a given number of future instances, are all questions of the utmost importance, which come within the scope of the calculus, and cannot, in fact, be justly appreciated by any other means.

The application of the theory of probability to the subject of jurisprudence, and the verdicts of juries and decisions of tribunals, has been discussed by the Marquis Condorcet in various articles in the *Encyclopédie Méthodique*; but more especially in his *Essai sur l'application de l'Analyse à la Probabilité des Décisions rendues à la Pluralité des Voix*, Paris 1785; a work of great ingenuity, and abounding with interesting remarks on subjects of the highest importance to humanity. James Bernoulli, it appears, had intended to treat jurisprudence as a branch of probability in the *Ars Conjectandi*, but his premature death prevented that work from being completed. There is a memoir on the subject by his nephew Nicolas, in the Leipsic Acts for 1711. The most important questions to be determined, are the number

of jurors of which a jury ought to consist, and the majority which should be required to agree in a verdict in order to afford, on the one hand, the greatest probability that an accused person will not be wrongly condemned; and, on the other, to give to society the greatest security that its interests will not be compromised, by allowing too great facilities for the guilty to escape. This important subject has been treated more profoundly, and with numerical elements derived from much better data than existed in the time of Condorcet, in a recent work by Poisson, to which we shall presently allude.

Another of the moral subjects to which the theory of probability has been applied, and connected with the preceding, is the appreciation of the evidence of testimony. In matters of this kind, it is easy to see that the calculus must be founded almost entirely upon hypothetical data. The veracity of a witness can scarcely be made the subject of direct experiment; and by reason of the complicated circumstances with which the facts forming the subject of testimony are usually accompanied, and the numberless ways in which mankind are influenced by their passions, credulity, or ignorance, it is perhaps equally impossible to deduce an average value from the comparison of a great number of statements which have been ascertained to be true or false. Numerical results can therefore only be obtained by having recourse to hypotheses, and consequently must be considered as only probable approximations. The knowledge, however, which is thus obtained of the various combinations of the quantities concerned, affords important aid in guiding our judgments in complicated cases, and when we have to decide upon conflicting testimony. Approximations deduced from a train of accurate and systematic reasoning, are al-

ways to be preferred to the most specious arguments drawn from any other source.

The analysis of probability has been applied with signal advantage in many researches of Natural Philosophy, but especially in appreciating the mean errors of observations. Owing to the imperfections of sense and of instruments, physical magnitudes are only susceptible of being measured within certain limits of accuracy; and where the last degree of precision is indispensable, as in practical astronomy, it is only by means of a very great number of measures, compared with one another, and combined according to the methods which this calculus points out, that we can obtain the nearest approximation to the true values which the observations are capable of giving. The mean errors of observations were treated as questions of probability by Lagrange in the *Turin Memoirs* for 1773; but it is to Laplace that the theory owes its principal extension and most important results. The method of combining numerous equations of condition now universally followed, known as the method of *minimum squares*, and which Laplace has demonstrated to be that which leaves the least probable amount of error in the final equations, was made known by Legendre in an Appendix to his *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*, published in 1806. A similar method, however, or rather the same, (for they are identical in principle,) had been discovered by Gauss, and employed by him for several years before the work of Legendre made its appearance.¹

Laplace's great work, the *Théorie Analytique des Probabilités*, first published in 1812, is one of the most remarkable productions that has ever appeared in abstract science. The

¹ *Theoria Motus Corporum Cælestium*, p. 221.

principles of the calculus, as well as the peculiar methods of analysis which it requires, and the most interesting and difficult questions which it presents, are here discussed in a far more general manner than had been attempted by any former writer on the subject ; and it may be said, accordingly, to have placed the theory under an entirely new aspect. It is much to be regretted that so little pains have been taken by the illustrious author to render the work intelligible to the generality of mathematical readers. Consisting for the greater part of separate memoirs presented at different times to the Academy of Sciences, arranged without regard to symmetry or order, it abounds with repetitions which only serve to embarrass the student ; while the deficiency of explanation combined with the subtlety of the analysis, and the inherent intricacy of the subject, render it often a painfully difficult task to seize the force of the demonstrations. Notwithstanding these defects, however, it forms one of the most splendid creations of mathematical genius ; and is alike admirable, whether we regard the extension which has been given to the calculus, or the results which have been arrived at, or the tone of lofty philosophy in which subjects bearing on some of the most important concerns of mankind are treated.

Next to the *Théorie Analytique* of Laplace, the most important work which has hitherto appeared on the subject of probability is the recent one of Poisson, entitled, *Recherches sur la Probabilité des Jugements*, (Paris 1837.) Although it might be inferred from the title that this work relates only to a single though very interesting application of the theory, the greater part of it is devoted to the development and demonstration of the general principles, and the discussion of the principal questions which present themselves in the different applications ; and it is only in the last of the five books

of which it consists that the special subject to which the title refers is taken into consideration. In applying the theory to the decisions of tribunals, Condorcet and Laplace had been unable to obtain positive results from the want of authentic data ; but the recent publication by the French government of the *Comptes Généraux de l'Administration de la Justice Criminelle*, in France, having furnished an immense collection of facts from which the requisite data could be obtained, Poisson was led to consider the subject anew, and the results of his investigations, which are of singular interest, are given in the work now mentioned. Poisson had already given a theory of the mean errors of observations in the *Additions to the Connaissance des Temps* for 1827 and 1832.

It is in these two works of Laplace and Poisson that the higher and more abstruse parts of the theory of probabilities must be studied. A very clear exposition of the principles, accompanied with many interesting remarks on the uses and applications of the theory, is given by Lacroix in his valuable little work, *Traité Élémentaire du Calcul des Probabilités*, Paris 1822.

Since the time of Demoivre, the English treatises on the general theory of probability have neither been numerous, nor, with one or two exceptions, very important. Simpson's *Laws of Chance* (1740) contains a considerable number of examples, in the solution of which the author displays his usual acuteness and originality, but as they belong entirely to that class in which the chances are known *a priori*, they give no idea of the most interesting applications of the theory. Dodson's *Mathematical Repository* contains a large selection of the same kind. The Essay in the *Library of Useful Knowledge*, by Mr. Lubbock, gives a more comprehensive and philosophical, though an elementary view of the sub-

ject ; but by far the most valuable work in the language is the Treatise in the *Encyclopedia Metropolitana*, by Professor De Morgan, 1837. In this very able production, Mr. De Morgan has treated the subject in its utmost generality, and embodied, within a moderate compass, the substance of the great work of Laplace.

Within the limits to which the present article must be confined, it would be hopeless to attempt giving a complete view of a branch of science which embraces so many complicated and intricate subjects of research, and which requires the aid of some of the most abstruse and recondite theories of the modern mathematics. In the higher applications of the theory, the analysis of many of the questions which arise, in order to be made intelligible, would require an extent of development and a parade of mathematical formulæ altogether incompatible with the plan and scope of this work. All that we can propose to ourselves, therefore, is to explain as briefly as may appear consistent with perspicuity, the general principles of the theory, and to give an outline of the manner in which these are applied to some of the more important questions which have been investigated by Laplace and Poisson. The examples will be selected with a view to shew the nature of the principal results of the mathematical theory, as well as the peculiar methods of analysis which are of most general application.

SECTION I.

GENERAL PRINCIPLES OF THE THEORY OF PROBABILITY.

1. The term *probable*, in its popular acceptation, is used in reference to any unknown or future event, to denote that in our judgment the event is more likely to be true than not, or more likely to happen than not to happen. Without attempting to make an accurate enumeration of the various circumstances which are favourable or unfavourable to its occurrence, or to balance their respective influences, we suppose there is a preponderance on one side, and accordingly pronounce it to be probable that the event has occurred, or will occur, or the contrary.

2. If we can see no reason why an event is more likely to happen than not to happen, we say it is a *chance* whether the event will happen or not; or if it may happen in more ways than one, and we have no reason for supposing it will happen in any one of these ways rather than in another, we say it is a *chance* whether it will happen in any assigned way or in any other. Suppose, for example, an unknown number of balls of different colours to be placed in an urn, from which a ball is about to be extracted by a person blindfold. Here we have no reason for supposing that the ball about to be drawn will be of one colour rather than another, that it will be white rather than black,

or red ; and accordingly say it is a *chance* whether the ball will come out of a particular colour, or a different. In this instance, then, the term *chance* denotes, simply, the absence of a known cause. If, however, we are made acquainted with the number of balls in the urn, and the number there are of each of the different colours, the term is used in a definite sense. For instance, suppose the urn to contain ten balls, of which nine are white, and the remaining one black, we say there are nine chances in favour of drawing a white ball, and one chance only in favour of drawing the black ball. Chance, in this sense, denotes a *way of happening*, or a particular *case* or *combination* that may arise out of a number of other possible cases or combinations ; and an event becomes probable or improbable according as the number of chances in its favour is greater or less than the number against it. *Chance* and *presumption* are also frequently used synonymously with *probability*.

3. The *mathematical probability* of any event is the ratio of the number of ways in which that event may happen to the whole number of ways in which it may either happen or fail. Thus, recurring to the previous example, the event, namely, the drawing of a ball from an urn containing 9 white balls and 1 black, may happen in 10 different ways, inasmuch as any one of the 10 balls may be drawn ; but in one only of those ways will the event be a black ball ; and therefore the probability of drawing the black ball is $\frac{1}{10}$. In like manner, as there are 9 different ways in which a white ball may be drawn, or 9 chances of drawing a white ball, and ten chances in all, the probability of drawing a white ball at the first trial is $\frac{9}{10}$. It follows immediately from this definition, that the probability of drawing a ball of either colour will remain the same,

however the number of balls in the urn may be increased, provided those of each colour are increased in the same proportion. For instance, suppose the number of white balls to be 45, and the number of black balls to be 5; the number of chances in favour of drawing a black ball is 5, while there are 50 chances in all, consequently the probability of a black ball being drawn is $\frac{5}{50} = \frac{1}{10}$. In the same manner, the probability of drawing a white ball is $\frac{45}{50} = \frac{9}{10}$; the same as before. Generally, let E and F be two contrary events, that is to say, such that the one or the other of them must necessarily happen, and both cannot happen together; and let a be the number of chances or combinations which produce the event E, and b be the number of combinations which produce the event F, or cause the failure of E; then the probability that E will happen is $\frac{a}{a+b}$; and the probability that F will happen, or that E will *not* happen is $\frac{b}{b+a}$. In future, the term probability will be used only to signify mathematical probability.

4. It is to be carefully remarked, that the different chances or combinations which form the elements of probability are supposed to be perfectly equal. If this equality does not hold, and there is any circumstance respecting the event under consideration which renders one combination or set of combinations more likely to occur than another, the different combinations must be multiplied by numbers proportional to their respective facilities, after which the units in each multiplier may be regarded as so many distinct chances, from which the probability of the event will be found by the above formula. This is equivalent to saying that a combination or chance which is twice

as likely to happen as another, must be regarded as two equal and similar combinations in comparison of that other ; a proposition which is sufficiently obvious.

5. It follows from the above definition, that the probability of any contingent event is measured by a fraction less than unity, and may have any value between 0 and 1. It follows, also, that the sum of the two fractions which measure the probabilities of two contrary events is equal to unit, which is the measure of certainty, inasmuch as either the one or the other necessarily occurs. Thus, in the last

example, the probability of the event E is $\frac{a}{a+b}$, and that of the contrary event F is $\frac{b}{a+b}$, and $\frac{a}{a+b} + \frac{b}{a+b} = 1$. Hence

if p denote the probability of any event E, and q the probability of the contrary event F, we have $q = 1 - p$. This consequence of the definition is of great importance in the calculation of probabilities.

6. We have here supposed the result of a trial to be necessarily one or other of two events E and F ; but it is easy to imagine the trial to be of such a kind that it may give rise to any one of a number of events E, F, G, H, &c. each having a given number of chances in its favour. This case is represented by supposing an urn to contain balls of as many different colours or sorts as there are different events. Let the urn be conceived to contain a balls of the sort which produces the event E, b of the sort which produces F, c of the sort which produces G, and so on ; and let $a + b + c + d$, &c. $= k$, so that k is the whole number of balls in the urn. The probabilities of the different events E, F, G, H, &c. are then, respectively, by the definition,

$$\frac{a}{k}, \frac{b}{k}, \frac{c}{k}, \frac{d}{k}, \&c.$$

the sum of which $=1$. In fact, if a ball be drawn at all, it must be of one or other of the different sorts contained in the urn; and consequently the sum of all the probabilities amounts to unit or certainty.

7. When an event is compounded of two or more simple events independent of each other, the probability of the compound event is equal to the product of the probabilities of the several simple events of which it is compounded. Let us imagine two urns, A and B, of which A contains a white balls and b black, and B contains a' white and b' black. Make $a+b=c$, and $a'+b'=c'$, and let the compound event whose probability is to be determined be the drawing of a white ball from both urns. Now, as each of the c balls in A may be drawn with any one of the c' balls in B, the whole number of ways in which the balls in A may be differently combined by pairs with the balls in B, or the whole number of possible cases is cc' . But the number of cases favourable to the compound event is evidently the number of different ways in which a white ball may be drawn from A with a white ball from B, and therefore equal to aa' . Hence by the definition (4), the probability that a white ball will be drawn from both urns is $\frac{aa'}{cc'}$. Now, if

p denote the probability of drawing a white ball from A, and p' that of drawing a white ball from B, we have by the definition $p = \frac{a}{c}$, and $p' = \frac{a'}{c'}$; whence $\frac{aa'}{cc'} = pp'$.

In general, let p denote the probability of an event E, p' that of another event E', p'' that of a third E'', and so on; then the probability of the concurrence of the events E,

E' , E'' , &c., or the probability that they will all happen, is $p \times p' \times p''$, &c.; that is to say, the probability of an event compounded of any number of simple and independent events, is the product of the respective probabilities of the several simple events.

The probabilities that the several simple events E , E' , E'' &c., will *not* all happen, or that some of them will happen and others fail, are easily determined in the same manner; it will be sufficient to indicate their several expressions. Suppose there are only three simple events, of which the probabilities are respectively p , p' , and p'' ; and let $q=1-p$, $q'=1-p'$, $q''=1-p''$. The product $pq'q''$ expresses the probability of the compound event which consists in E happening and E' and E'' both failing; $qp'q''$ is the probability that E' will happen, and that E and E'' will both fail; $pp'p''$ is the probability they will all three happen; $1-pp'p''$ is the probability they will *not* all three happen, or that one of them *at least* will fail; $qq'q''$ is the probability they will all fail; and $1-qq'q''$ is the probability they will *not* all three fail, or that one *at least* of them will happen.

8. As an example of the application of this rule, suppose it were required to assign the probability of throwing aces, at one throw, with two common dice. As a common die has six symmetrical faces, there are in respect of each die six ways equally possible, in which the simple event may happen. The probability therefore of throwing ace with one die is $\frac{1}{6}$, that is, $p=\frac{1}{6}$. In respect of the second die, we have also $p'=\frac{1}{6}$; hence the probability of the compound event, or that aces will be thrown is $pp'=\frac{1}{6} \times \frac{1}{6}=\frac{1}{36}$. The probability that aces will *not* be thrown at any assigned trial is therefore (5) $1-\frac{1}{36}=\frac{35}{36}$; and the *odds* against throwing aces at any given trial are 35 to 1.

Again, suppose two numbers, each consisting of 7 digits, to be taken at random, (for instance from a table of logarithms), and let it be proposed to assign the probability that the subtraction of the one from the other will be performed without its being necessary, in any case, to increase the upper figure. Here, as each digit may have any one of the ten values from 0 to 9 both inclusive, and as each of those values in the upper line may be combined with any one of them in the lower line, there are 100 different combinations or equally possible cases for each partial subtraction. Now, if the upper figure be 0, there is only one of those cases favourable to the event, or which will admit of the subtraction being performed, namely, when the figure below is also 0. If the upper figure be 1, there are two cases favourable, namely, those in which the under figure is 0 or 1. If the upper figure be 2, there are three favourable cases, namely, when the under figure is 0, 1, or 2. Proceeding in this way through all the digits, the whole number of favourable cases is found to be

$$1+2+3+4+5+6+7+8+9+10=55.$$

Hence, for each partial subtraction there are 55 favourable cases out of 100 possible cases; therefore (4) the probability that any one of the figures in the upper line is not less than the corresponding figure in the under line is $\frac{55}{100}$; and we have $p=p'=p''=\&c.=\frac{55}{100}$ for the probability of each of the seven simple events or partial subtractions, whence, by (7), the probability of the compound event is

$$p \times p' \times p'' \times \&c. = \left(\frac{55}{100} \right)^7 = (.55)^7 = .0152243,$$

which is less than $\frac{1}{65}$, and greater than $\frac{1}{68}$.

9. When an event may happen in several different ways,

each independent of the others, the probability of the event is the sum of all the partial probabilities taken in respect of each of the different ways.

Suppose there are n different urns $A_1, A_2, A_3, \dots, A_n$, each containing balls of two colours, white and black, and let the whole number of balls in each urn respectively, be

$$c_1, c_2, c_3, \dots, c_n,$$

and the number of white balls in each be

$$a_1, a_2, a_3, \dots, a_n,$$

and let the event E be the extraction of a white ball in drawing a ball from any urn at random. In this case there are n different ways, all equally probable, in which the event may happen, for it may be drawn with equal facility from any one of the urns. The probability that the ball will be

drawn from any given urn, A_1 , is therefore $\frac{1}{n}$; and if it

be drawn from this urn, the probability of its being white is $\frac{a_1}{c_1}$; therefore, by (7), the probability of a white ball being

drawn from A_1 is $\frac{1}{n} \cdot \frac{a_1}{c_1}$. In like manner the probability of a white ball being drawn from A_2 is shewn to be

$\frac{1}{n} \cdot \frac{a_2}{c_2}$; from A_3 to be $\frac{1}{n} \cdot \frac{a_3}{c_3}$, and so on. Denoting

therefore by p the whole probability of the event E , the proposition affirms that

$$p = \frac{1}{n} \left(\frac{a_1}{c_1} + \frac{a_2}{c_2} + \frac{a_3}{c_3} + \dots + \frac{a_n}{c_n} \right).$$

To prove this, let the fractions $\frac{a_1}{c_1}, \frac{a_2}{c_2}$ &c. be reduced to a common denominator, and suppose the equivalent fractions to be

$$\begin{aligned} \text{prob of drawing the ball in } A_1 &= \frac{a_1}{c_1} \left\{ \frac{c_2}{c_1 c_2} - \frac{c_3}{c_1 c_3} - \dots \right\} \\ &= \frac{a_1 c_2}{c_1 c_2} - \frac{a_1 c_3}{c_1 c_3} + \dots \\ \therefore \text{prob of drawing the ball in } A_2 &= \frac{a_2}{c_2} \left\{ \frac{c_1}{c_1 c_2} - \frac{c_3}{c_2 c_3} - \dots \right\} \\ &= \frac{a_2 c_1}{c_1 c_2} - \frac{a_2 c_3}{c_2 c_3} + \dots \end{aligned}$$

$$\frac{a_1}{\gamma}, \frac{a_2}{\gamma}, \frac{a_3}{\gamma}, \dots \frac{a_n}{\gamma}.$$

We may now conceive the urns $A_1, A_2, A_3 \dots A_n$ to be replaced by others, each containing the same number, γ , of balls, and of which the first contains a_1 white balls, the second a_2 , and so on; and it is evident that the chance of a white ball being drawn from this new system of urns will be precisely the same as it was for a white ball being drawn from the first system. Now the probability of drawing a white ball from the new system will not be altered by placing the whole of the $n\gamma$ balls in a single urn, for they may still be conceived as arranged in groups, disposed in any manner whatever, each group containing the same number of balls, and the same proportion of white to black as were in the separate urns; and as each group contains the same number of balls, the chance of laying the hand on any one group is the same as that of laying it on any other. The probability of drawing a white ball from the single urn, is therefore the same as for drawing it from the group of separate urns which contain each the same number of balls. But the probability of drawing it from the single urn is the ratio of the number of white balls contained in the urn to the number of both colours, therefore (this probability being p) we have

$$p = \frac{1}{n\gamma} (a_1 + a_2 + a_3 \dots + a_n);$$

whence, substituting for $\frac{a_1}{\gamma}, \frac{a_2}{\gamma}, \&c.$, their respective

values, $\frac{a_1}{c_1}, \frac{a_2}{c_2}, \&c.$, we have

$$p = \frac{1}{n} \left(\frac{a_1}{c_1} + \frac{a_2}{c_2} + \frac{a_3}{c_3} \dots + \frac{a_n}{c_n} \right).$$

As a particular case suppose three urns A, B, C to be placed

together, of which A contains 2 white balls and 1 black ; B 3 white balls and 2 black, and C 4 white and 3 black, and let it be required to determine the probability p of a white ball being drawn from the group by a person who is ignorant of the contents of the different urns. As there is no reason for selecting one urn in preference to another, the probability that he will put his hand into the urn A is $\frac{1}{3}$; and if he draw from this urn the probability that a white ball will be drawn is $\frac{2}{3}$, there being 2 cases favourable to that event, and 3 cases in all. The probability of both events is therefore $\frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$. In like manner, the probability of the ball being drawn from B is $\frac{1}{3}$; and if drawn from B the probability of its being white is $\frac{3}{5}$; therefore, the probability of this compound event is $\frac{1}{3} \times \frac{3}{5} = \frac{1}{5}$. Lastly, the probability of the ball being drawn from C is $\frac{1}{3}$; and if drawn from C the probability of its being white is $\frac{4}{7}$; therefore, the probability of this compound event is $\frac{1}{3} \times \frac{4}{7} = \frac{4}{21}$. Hence, by the proposition now demonstrated, the complete probability of the event E is

$$p = \frac{2}{9} + \frac{1}{5} + \frac{4}{21} = \frac{19}{45}.$$

If all the balls had been placed in a single urn, the probability of drawing a white ball would have been $\frac{9}{15}$, for there are $3 + 5 + 7 = 15$ balls in all, of which $2 + 3 + 4 = 9$ are white. But $\frac{9}{15} = \frac{3}{5}$; a fraction which differs sensibly from $\frac{19}{45}$, the measure of the probability of the same event when the balls are distributed in the manner above supposed amongst the different urns. The distinction between the two cases is important.

10. The rule laid down in (7) for finding the probability of a compound event applies alike whether the simple events are determined simultaneously or in succession. In fact, when the simple events are entirely independent of each

other, the chances which determine the compound event are not influenced in any way by the intervention of time. Suppose, for example, the compound event to be the throwing of a certain number of points with a given number m of dice; the chances for and against the event are obviously the same whether the m dice are thrown at once, or a single die is thrown m times successively. But as the determination of the probability of a compound event is in general facilitated by supposing the simple events to be decided one after the other, it will be convenient to view the subject in this light in explaining the method of forming the different combinations of the chances by which the probabilities of compound events are determined.

SECTION II.

OF THE PROBABILITY OF EVENTS DEPENDING ON A REPETITION OF TRIALS, OR COMPOUNDED OF ANY NUMBER OF SIMPLE EVENTS, THE CHANCES IN RESPECT OF WHICH ARE KNOWN *A PRIORI*, AND CONSTANT.

11. Suppose an urn to contain $a + b$ balls, a white and b black, and let a ball be successively drawn, and replaced in the urn after each drawing, in order that the chances in favour of drawing a ball of either colour may be the same in every trial, and let it be required to find the respective probabilities of the different possible results of any number of drawings.

Let us first suppose the number of trials to be two. The event may happen in any of these four different ways: first white, second white; first white, second black; first black, second white; first black, second black. Assuming W to represent the simple event which consists in the drawing of a white ball, and B that of a black ball, and supposing the order of the arrangement of the two letters to correspond with the order of succession of the simple events, the four possible cases or combinations will be represented thus:—

$WW, WB, BW, BB.$

Now let the probability of drawing a white ball in any trial be p , and that of drawing a black ball be q , (whence,

$p = \frac{a}{a+b}, q = \frac{b}{a+b}$) the probabilities of the four possible compound events are by (7) respectively as under :

$$\begin{aligned}\text{probability of WW} &= p \times p = p^2 \\ \text{of WB} &= p \times q = pq \\ \text{of BW} &= q \times p = pq \\ \text{of BB} &= q \times q = q^2\end{aligned}$$

If we disregard the order of succession, and consider the two arrangements WB and BW, which are equally probable, as forming the same compound event, namely, a ball of both colours in the two trials, the probability of this event, by (9), becomes $2pq$. The sum of the probabilities of all the possible arrangements is therefore

$$p^2 + 2pq + q^2 = (p + q)^2;$$

whence it appears that the probabilities of the different arrangements in two trials are respectively the terms of the development of the binomial, $(p + q)^2$.

Let us next suppose the number of trials to be three. The different arrangements that may be formed of the simple events in three trials, with the probability of each respectively, are as follows :—

$$\begin{aligned}\text{WWW, probability of which} &= ppp = p^3 \\ \text{WWB,} &= ppq = p^2q \\ \text{WBW,} &= pqp = p^2q \\ \text{BWW,} &= qpp = p^2q \\ \text{WBB,} &= pqq = pq^2 \\ \text{BWB,} &= qpq = pq^2 \\ \text{BBW,} &= qq p = pq^2 \\ \text{BBB,} &= qqq = q^3\end{aligned}$$

It thus appears that the probability of obtaining two events of one kind, and one of the other, is the same in whatever order they succeed each other, and, in fact, is independent

of the order. Disregarding, then, the order of succession, and considering the combination of two white balls with one black, in whatever order they may be arranged, as the same compound event, the probability of its occurrence *in any order whatever*, being the sum of its probabilities in each particular order (9), is $3p^2q$. In like manner, regarding the combination of two black balls with one white, in any order of arrangement, as the same compound event, its probability is $3pq^2$. The compound event resulting from three trials must then happen in one of four different ways, namely, 3 white balls; 2 white, combined with 1 black, in any order; 2 black, combined with one white, in any order; or, lastly, 3 black; and the sum of the probabilities of these different cases is

$$p^3 + 3p^2q + 3pq^2 + q^3 = (p + q)^3.$$

Hence the probabilities of all the different possible combinations in three trials are respectively given by the development of the binomial $(p + q)^3$.

12. In general, let p denote the probability of any simple event E, then the probability of E happening twice in two trials is p^2 , of happening thrice in 3 trials p^3 , and of happening m times in m successive trials, p^m . In like manner, the probability of the contrary event F being q ($p + q = 1$), the probability of F happening n times in n successive trials is q^n . Hence (7) the probability of E happening m times, and then F happening n times in succession, in $m + n$ trials, is $p^m q^n$. But the probability of these events happening in any assigned order is the same as that of their happening in any other assigned order; therefore $p^m q^n$ is the measure of the probability that E will occur m times, and F will occur n times in a determinate order. Now, let $m + n = h$, and let U be the number of different ways in which m events E, and

n events F, can be combined in h trials, and P be the probability of any one of these combinations whatever, or the probability of E occurring m times, and F occurring n times in h trials, without regard to the order in which they succeed each other, we have then

$$P = Up^mq^n.$$

In order to determine the value of U, we may suppose the events in question to be so many different things represented by the letters A, B, C, D, E, &c. of which there are m of one kind, and n of another, and make $m+n=h$; then by the algebraic theory of combinations, we have

$$U = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n}.$$

This value of U is symmetrical in respect of m and n , and may be otherwise written in either of the two following forms,

$$U = \frac{h(h-1)(h-2) \dots h-m+1}{1 \cdot 2 \cdot 3 \dots m},$$

$$U = \frac{h(h-1)(h-2) \dots h-n+1}{1 \cdot 2 \cdot 3 \dots n},$$

which shew that the probability P, or the product Up^mq^n is the $(m+1)$ th term of the development of the binomial $(p+q)^h$ arranged according to the increasing powers of p , or the $(n+1)$ th term of the same development arranged according to the increasing powers of q . Hence we conclude that when p and q remain constant, the probabilities of all the different compound events which can be formed by the combination of the simple events E and F in h trials, are expressed by the different terms of the formula $(p+q)^h$ expanded by the binomial theorem.

The whole number of possible cases is evidently $h+1$, for in h experiments, E may occur h times, $h-1$ times,

$h-2$ times..... $h-h$ times ; this last being the case in which the contrary event F occurs in all the trials. The different cases are unequally probable, both by reason of the greater or smaller number of combinations by which they may be produced, and which in reference to each case is represented by U, and by reason of the inequality between p and q . It will be shewn afterwards, that when $p=q$, and h is a whole number, the most probable case is that in which the occurrences of E and F are equal ; and if h is an odd number, the two most probable cases are those in which the difference in the number of occurrences of E and the number of occurrences of F is unity.

13. In order to place the proposition now demonstrated in a clearer light, let us consider separately the different terms of the development of $(p+q)^h$, namely,

$$\begin{aligned}
 & p^h + hp^{h-1}q + \frac{h(h-1)}{1 \cdot 2} p^{h-2}q^2 \dots\dots\dots \\
 & + \frac{h(h-1)(h-2)\dots\dots h-n+1}{1 \cdot 2 \cdot 3 \dots\dots n} p^{h-n}q^n \\
 & \dots\dots\dots + q^h.
 \end{aligned}$$

The first term p^h expresses the probability that the event E will in every one of the h trials. The second term $hp^{h-1}q$ expresses the probability that E will occur $h-1$ times, and F once, without distinction of order ; that is to say F may happen at the first or last or any intermediate trial. If a determinate succession is proposed, for example, that of $h-1$ times the event E in succession, and F in the next trial, the probability of the event in the assigned order is found by suppressing the coefficient h , and is consequently $p^{h-1}q$.

The third term $\frac{h(h-1)}{1 \cdot 2} p^{h-2}q^2$ expresses the probability

that the result of h trials will be $h-2$ times the event E, and twice the event F, without distinction of order. If a particular order be assigned, it is necessary to suppress the coefficient, and the probability of the simple events occurring in that particular order is $p^{h-2}q^2$.

The general term $\frac{h(h-1)(h-2)\dots\dots h-n+1}{1 \cdot 2 \cdot 3 \dots\dots n} p^{h-n} q^n$ expresses the probability that the result of h trials will be $(h-n)$ times the event E, and n times the event F in any order. The probability of $(h-n)$ times E and n times F in an assigned order is $p^{h-n}q^n$.

14. If we suppose the event E to be such that the chances in favour of its happening or failing are equal, that is, if $p=q=\frac{1}{2}$, the different terms of the binomial $(p+q)^h$, on suppressing the coefficients, become all equal; so that a particular order being assigned in each of the possible cases or combinations, all the cases become equally probable. Thus, suppose a shilling to be tossed 100 times in succession, the probability of head turning up in every trial is $(\frac{1}{2})^{100}$. The probability of 50 heads and 50 tails in any assigned order is $(\frac{1}{2})^{50} \times (\frac{1}{2})^{50} = (\frac{1}{2})^{100}$; if $m+n=100$, the probability of m heads and n tails is also $(\frac{1}{2})^m (\frac{1}{2})^n = (\frac{1}{2})^{m+n} = (\frac{1}{2})^{100}$. Hence the probability of any compound event formed by the combination of two simple contrary events succeeding each other in an assigned order, and each having the same probability, is independent of the ratio of the simple events, and depends only on the number of trials. Before the trials, it is an even wager that head will be turned up in succession 100 times, and that the result of 100 trials will be 50 heads and 50 tails in a given order of succession, or any proportion of heads to tails in an order arbitrarily chosen. This consideration is frequently lost sight of in reasoning about those

events of the natural world, which are termed extraordinary and miraculous. If in tossing a shilling 100 times into the air, the number of heads turned up is found nearly equal to the number of tails, the event excites no surprise ; something like it was expected. On the contrary, if the difference between the number of heads and the number of tails is considerable, the event is termed extraordinary ; and if head turned up in every trial without exception, we should scarcely be persuaded that such an event was entirely the result of chance, and independent of a special cause. Nevertheless, the *a priori* probability that every trial will give head, is precisely the same as the probability of throwing any given number of heads and tails in an assigned order of succession. It will, however, be proved afterwards, that if such an event as throwing head 100 times in succession were actually observed, the probability of a special cause having intervened, would approach very nearly to certainty.

15. Hitherto we have supposed the compound event to be formed by the combination of two simple events only, E and F, one of which necessarily excludes the other. Let us now suppose there are any number of simple events, E_1 , E_2 , E_3 , &c. of which the respective probabilities are p_1 , p_2 , p_3 , &c. and such that one or other of them necessarily happens in each trial, so that $p_1 + p_2 + p_3 + \dots = 1$, and determine the probability of any assigned combination of them in a given number of trials. This case may be represented by supposing an urn to contain a number of balls of as many different colours as there are distinct events ; the event E_i will be the drawing of a ball of the colour i , and its probability p_i will be the fraction whose numerator is equal to the number of balls of the colour i , and denominator the whole number of balls in the urn. Now the probability of the

event E_1 happening m times in succession is p_1^m by (12); that of E_2 happening n times in succession is p_2^n ; that of E_3 happening r times in succession p_3^r ; and so on. Therefore (7) the probability of the compound event which is formed by the occurrence of m times E_1 , n times E_2 , r times E_3 , and so on, these events succeeding each other in order, is the product $p_1^m p_2^n p_3^r$, &c. But the probability of the simple events succeeding each other in any particular order is the same as that of their succeeding in any other assigned order (12); consequently, if U' denote the number of different ways in which m events E_1 , n events E_2 , r events E_3 , &c. can be combined, or succeed each other, and P' be the probability of the compound event in any order whatever, we have,

$$P' = U' p_1^m p_2^n p_3^r, \&c.$$

Assuming $h = m + n + r + \&c.$ we have also by the theory of combinations,

$$U' = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n \times 1 \cdot 2 \cdot 3 \dots r \times \&c.}$$

the factor U' being the coefficient of the term which has for its multiplier $p_1^m p_2^n p_3^r$, &c. in the expansion of the multinomial $(p_1 + p_2 + p_3 + \&c.)^h$, whence

$$P' = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n \times 1 \cdot 2 \cdot 3 \dots r \times \&c.} p_1^m p_2^n p_3^r, \&c.$$

We shall now proceed to give some examples of the applications of the preceding formulæ.

16. Let it be proposed to assign the probability P , of throwing ace *once*, and not oftener, in four successive throws of the same die.—Simpson, p. 15.

Here, the chance of throwing ace in a single trial being $\frac{1}{6}$, we have $p = \frac{1}{6}$, and consequently $q = \frac{5}{6}$, and also $h = 4$. Now the compound event being the occurrence of the sim-

ple event E, whose probability is p , once, and of the contrary event F three times, the probability of the compound event is that term of the development of $(p+q)^4$ which is multiplied by pq^3 . If, therefore, in the formula,

$$P = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n} p^m q^n,$$

we make $p = \frac{1}{6}$, $q = \frac{5}{6}$, $h = 4$, $m = 1$, $n = 3$, we shall have

$$P = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \times 1 \cdot 2 \cdot 3} \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 = \frac{125}{324},$$

which is the probability required, and the same as that of throwing one ace, and not more than one, at a single throw with 4 dice.

The probability of the contrary event, that is to say, the probability of either not throwing an ace at all, or of throwing more aces than one is $1 - \frac{1}{3} \frac{2}{2} \frac{5}{4} = \frac{1}{3} \frac{9}{2} \frac{9}{4}$; and therefore the odds against throwing one ace and no more in 4 throws of a common die are 199 to 125, or 8 to 5 very nearly.

17. If in this example it had been proposed to assign the probability of throwing ace once *at least*, instead of once and not more, it would have been necessary to have included those cases in which the ace occurs twice, or three times, or in each of the four trials. The binomial $(p+q)^4$ gives

$$p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4,$$

the first term of which expresses the probability of throwing ace four times in succession; the second that of throwing ace three times, and another number once; the third that of throwing ace twice, and a different face twice; the fourth that of throwing ace once, and a different face three times; and the fifth that of throwing a different face in each of the four trials. But as every one of these compound events, excepting the last, satisfies the condition of ace being thrown once at least, the whole probability of that event

must be the sum of the probabilities of the different events by which it may be produced (9) and is consequently

$$\left(\frac{1}{6}\right)^4 + 4\left(\frac{1}{6}\right)^3 \frac{5}{6} + 6\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 + 4\left(\frac{1}{6}\right) \left(\frac{5}{6}\right)^3 = \frac{671}{1296}$$

In general, the sum of the first $n+1$ terms of $(p+q)^h$ expresses the probability of obtaining *not less* than $h-n$ events, the probability of each of which is p , or *not more* than n contrary events, the probability of each of which is q .

Since $p+q=1$, the sum of all the terms of the series produced by the expansion of $(p+q)^h$ is equal to unit, and therefore the sum of any number of the terms is equal to unit diminished by the sum of the remaining terms. This consideration frequently gives the means of abridging the calculations. Thus, in the preceding example, instead of expanding the binomial $\left(\frac{1}{6} + \frac{5}{6}\right)^4$ in order to find the probabilities of throwing 4 aces, 3 aces, 2 aces, and 1 ace only, in a series of 4 trials, we might have sought the probability of not throwing ace at all. The probability of not throwing ace in a single trial is $\frac{5}{6}$, and therefore (7) that of not throwing it in 4 trials is $\left(\frac{5}{6}\right)^4 = \frac{625}{1296}$. Hence the probability of the contrary event, namely, that ace will be thrown once or oftener, is $1 - \frac{625}{1296} = \frac{671}{1296}$; the same as before.

18. Let a shilling be tossed; what is the probability that more than 3 heads will turn up in the first 10 trials? In this case, $p = \frac{1}{2}$, $q = \frac{1}{2}$, $h = 10$; therefore $(p+q)^h = \left(\frac{1}{2} + \frac{1}{2}\right)^{10} = \left(\frac{1}{2}\right)^{10} (1+1)^{10}$. Now the last term of this development expresses the probability that head will not turn up in any one of the ten trials; the last but one, the probability that it will turn up once; the last but two, the probability that it will turn up twice; and the last but three, the probability that it will turn up three times; therefore the four last

terms include all the different ways in which the ten trials give not more than three heads; and their sum consequently expresses the probability that *not more* than 3 heads will be thrown. Now the last four (or first four) terms of the expansion of $(1+1)^{10}$ are

$$1, 10, \frac{10 \cdot 9}{1 \cdot 2}, \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3},$$

and their sum is 176, which multiplied by $(\frac{1}{2})^{10} = \frac{1}{1024}$, gives $\frac{176}{1024}$, for the probability that not more than 3 heads will turn up; whence the probability of the contrary event or that more than 3 heads will be thrown, is $1 - \frac{176}{1024} = \frac{848}{1024} = \frac{53}{64}$; and the odds in favour of throwing heads more than three times in 10 trials are 53 to 11.

19. A and B engage in play; the probability of A's winning a game is p , and the probability of B's winning a game is q ; required the probability P , of A's winning m games before B wins n games, the play being supposed to terminate when either of those events has occurred.

It is evident that the question must be decided *at the latest*, by the $(m+n-1)$ th game; for supposing $m+n-2$ games to have been played, there is only one combination according to which the match can remain undecided, namely, that in which A has won $m-1$, and B $n-1$ games; and in this case the next game necessarily decides the match.

Suppose $m+x$ games to have been played. The probability that of these games m have been won by A, and x by B, is represented by the term of the binomial $(p+q)^{m+x}$ in which the factor $p^m q^x$ occurs (13); which term is

$$\frac{1 \cdot 2 \cdot 3 \dots m+x}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots x} p^m q^x.$$

But A cannot win m games out of $m+x$ exactly unless he wins the last game, for otherwise he must have won m games

out of $m+x-1$, if not out of a smaller number. In order therefore that A may win m games out of $m+x$ exactly, it is necessary in the first place that he wins $m-1$ out of $m+x-1$ in any order, and then that he wins also the next game. Now the probability of his winning $m-1$ games out of $m+x-1$ in any order (13) is

$$\frac{1 \cdot 2 \cdot 3 \dots m+x-1}{1 \cdot 2 \cdot 3 \dots m-1 \times 1 \cdot 2 \cdot 3 \dots x} p^{m-1} q^x;$$

and the probability of his winning the following game is p , whence the probability of both events is (7)

$$\frac{1 \cdot 2 \cdot 3 \dots m+x-1}{1 \cdot 2 \cdot 3 \dots m-1 \times 1 \cdot 2 \cdot 3 \dots x} p^m q^x,$$

which, therefore, expresses the probability of A's winning m games out of $m+x$ exactly.

If we suppose $x=0$, this formula becomes p^m , which is the probability of A's winning m games in succession. If $x=1$, it becomes mp^mq , the probability that A wins m games out of $m+1$. If $x=2$, it becomes $\frac{m(m+1)}{1 \cdot 2} p^m q^2$, the probability that A wins m games out of $m+2$. If $x=3$, it becomes $\frac{m(m+1)(m+2)}{1 \cdot 2 \cdot 3} p^m q^3$, the probability that A wins m games out of $m+3$; and so on. Continuing this process till we arrive at the term multiplied by $p^m q^x$, the sum of the probabilities of all the different compound events is

$$p^m \left\{ 1 + mq + \frac{m(m+1)}{1 \cdot 2} q^2 \dots + \frac{m(m+1) \dots m+x-1}{1 \cdot 2 \dots x} q^x \right\};$$

which expresses the probability of A's winning m games out of a number *not greater* than $m+x$.

Now it has been shewn, that the match is necessarily decided by $(m+n-1)$ games; consequently the solution

of the question is obtained by substituting $n-1$ for x in the last formula, which will then express the probability of A's winning m games in any order, out of a number not greater than $m+n-1$. On making this substitution, we obtain

$$P = p^m \left\{ 1 + mq + \frac{m(m+1)}{1 \cdot 2} q^2 + \dots + \frac{m(m+1) \dots m+n-2}{1 \cdot 2 \dots n-1} q^{n-1} \right\}.$$

The probability Q that the match will be decided in favour of B, or that B will win n games out of a number not greater than $m+n-1$, is found by changing m into n , and p into q , and is therefore

$$Q = q^n \left\{ 1 + np + \frac{n(n+1)}{1 \cdot 2} p^2 + \dots + \frac{n(n+1) \dots m+n-2}{1 \cdot 2 \dots m-1} p^{m-1} \right\}.$$

As an example, let us suppose $p = \frac{2}{3}$, $q = \frac{1}{3}$, $m=4$, and $n=2$. The probability of A's winning the match, or the value of P, becomes

$$\left(\frac{2}{3}\right)^4 \left\{ 1 + 4 \cdot \frac{1}{3} \right\} = \frac{112}{243};$$

and the probability of B's winning the match, or the value of Q,

$$\left(\frac{1}{3}\right)^2 \left\{ 1 + \frac{4}{3} + \frac{2 \cdot 3}{1 \cdot 2} \left(\frac{2}{3}\right)^2 + \frac{2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3} \left(\frac{2}{3}\right)^3 \right\} = \frac{131}{243}.$$

In this example the skill of A is supposed to be twice as great as that of B, and the number of games that must be won by him in order to gain the match is also twice as great as the number required to be won by B in order that B may

gain ; one might therefore suppose, that when they begin to play the chances in favour of each are equal. But the result shews that the chances in favour of A are fewer than those in favour of B in the proportion of 112 to 131 ; whence it appears that it would be unsafe to wager that a player who has two chances in his favour while his adversary has only one, will gain four games before his adversary shall have gained two.

Suppose A and B, engaged in play, agree to leave off before the match is decided, it is evident that the stakes ought to be shared between them in proportion to their respective probabilities of winning, and consequently the share of each is found from either of the above expressions for P and Q. This was one of the questions proposed by the Chevalier de Méré to the celebrated Pascal, to which allusion has already been made.

20. An urn contains $n+1$ balls, marked with the numbers 0, 1, 2, 3..... n ; a ball is successively drawn and replaced in the urn, so that the chance of drawing any given number remains the same in each trial, what is the probability that in h trials the sum of the numbers drawn will be equal to s ?¹

The solution of this problem depends on the number of ways in which the number s can be formed by the addition of h different numbers, each of which may have any value from 0 to n . If we suppose the numbers marked on the balls to be indexes of a certain quantity x , and develop the expression $(x^0 + x^1 + x^2 + \dots + x^n)^h$, the coefficient of any term of the development will indicate the number of different ways in which the balls may be drawn, so that the sum of the numbers drawn in h trials shall be equal to the sum of the in-

¹ Demoivre, *Miscellanea Analytica*, p. 196 ; Laplace, *Essai sur les Probabilités*, p. 253, et seq.

dexes of x in that term. If, therefore, we denote by N the coefficient of that term of the development in which the sum of the indexes is s , then N will be the number of cases favourable to the event. But the whole number of possible cases is $(n+1)^h$; therefore the probability of the event is $N \div (n+1)^h$.

On account of the particular form of the polynomial in question, the value of N is found without difficulty.

Because $x^0 + x^1 + x^2 + \dots + x^n = \frac{1-x^{n+1}}{1-x}$, therefore

$(x^0 + x^1 + x^2 + \dots + x^n)^h = (1-x^{n+1})^h (1-x)^{-h}$. Now, expressing these two factors in series, we have $(1-x^{n+1})^h$

$$= 1 - hx^{n+1} + \frac{h(h-1)}{1 \cdot 2} x^{2(n+1)} - \frac{h(h-1)(h-2)}{1 \cdot 2 \cdot 3} x^{3(n+1)} + \&c.$$

$$(1-x)^{-h} = 1 + hx + \frac{h(h+1)}{1 \cdot 2} x^2 + \frac{h(h+1)(h+2)}{1 \cdot 2 \cdot 3} x^3 + \&c.;$$

and the coefficients of the several terms of the product of these two series in which the sum of the indexes is s will be found as follows:—

(1.) Multiply the first term of the first series by that term of the second series of which the argument is x^s ; the coefficient of the product will be $\frac{h(h+1)(h+2) \dots h+s-1}{1 \cdot 2 \cdot 3 \dots s}$.

(2.) Multiply the second term of the first series by that term of the second series which has for its argument x^{s-n-1} ; the coefficient of the product will be

$$-h \times \frac{h(h+1)(h+2) \dots h+s-n-2}{1 \cdot 2 \cdot 3 \dots s-n-1}.$$

(3.) Multiply the third term of the first series by that term of the second series which has for its argument $x^{s-2(n+1)}$; the coefficient of the product will be

$$\frac{h(h-1)}{1.2} \times \frac{h(h+1)(h+2).....h+s-2n-3}{1.2.3 \quad \dots \quad s-2n-2}.$$

(4.) Proceed in the same manner with the fourth term of the first series, and so on with the others, advancing at each new multiplication one term to the right in the first series, and $n+1$ terms to the left in the second series, until a term is reached in the first series, the exponent of x in which is equal to, or greater than s . The sum of the several products thus obtained will be the value of N . We have therefore

$$\begin{aligned} N = & \frac{h(h+1)(h+2).....h+s-1}{1.2.3 \quad \dots \quad s} \\ & + \frac{h}{1} \times \frac{h(h+1)(h+2).....h+s-n-2}{1.2.3 \quad \dots \quad s-n-1} \\ & + \frac{h(h-1)}{1.2} \times \frac{h(h+1)(h+2).....h+s-2n-3}{1.2.3 \quad \dots \quad s-2n-2} \\ & - \&c. \end{aligned}$$

The series now found for N may be changed into another, having a more elegant form, by reducing all the terms to others having the common denominator $1.2.3.....h-1$. This will be accomplished by leaving out of the numerator and denominator of the first term all the numbers after $h-1$ to s , (including s), when s is greater than $h-1$, or by inserting the numbers between s and $h-1$ (the last included), when $s \leq h-1$; by leaving out of the numerator and denominator of the second term all the numbers from $h-1$ to $s-n$, or by inserting those numbers; and so on with the other terms. If we then make the common denominator $1.2.3...h-1=k$, we shall have

$$\begin{aligned} N = & \frac{1}{k} (s+1)(s+2)(s+3).....(s+h-1) \\ & - \frac{h}{k} (s-n)(s-n+1).....(s-n+h-2) \end{aligned}$$

$$+ \frac{h(h-1)}{1 \cdot 2k} (s-2n-1)(s-2n) \dots (s-2n+h-3) \\ - \&c.$$

to be continued till the last factor of one of the terms becomes 0 or negative. If we also make

$$s+h-1=f \\ s-n+h-2=f-(n+1)=f' \\ s-2n+h-3=f-2(n+1)=f'' \\ s-3n+h-4=f-3(n+1)=f''' \\ \&c.$$

and write the factors in each of the terms in the reverse order, the above value of N will become

$$N = +f(f-1)(f-2) \dots (f-h+2) \frac{1}{h} \\ -f'(f'-1)(f'-2) \dots (f'-h+2) \frac{h}{k} \\ +f''(f''-1)(f''-2) \dots (f''-h+2) \frac{h(h-1)}{1 \cdot 2k} \\ -f'''(f'''-1)(f'''-2) \dots (f'''-h+2) \frac{h(h-1)(h-2)}{1 \cdot 2 \cdot 3k} \\ + \&c.$$

21. As an example of the application of this formula, let it be required to assign the probability of throwing the point 16 with 4 common dice. (Simpson, p. 53.)

A die having no face marked 0, it is necessary, in order to adopt the formula to this case, to suppose the number of points on each face to be diminished by unit, which is equivalent to supposing $s-h$ to be substituted for s . The numbers are then 0, 1, 2, 3, 4, 5, and we have $n=5$, $h=4$, and $s=(16-4)=12$. Hence

$$f=s+h-1=15 \\ f'=f-(n+1)=9$$

$$f'' = f - 2(n + 1) = 3$$

$$f''' = f - 3(n + 1) = -3,$$

and $k=1.2.3$. Substituting these values in the formula, we find

$$N = 15 \cdot 14 \cdot 13 \times \frac{1}{6} (= +455)$$

$$-9 \cdot 8 \cdot 7 \times \frac{4}{6} (= -336)$$

$$+3 \cdot 2 \cdot 1 \times \frac{4 \cdot 3}{1 \cdot 2 \cdot 6} (= +6)$$

or $N=125$. Now the probability of the event is $N \div (n+1)^h$; and in the present case $(n+1)^h = 6^4 = 1296$; consequently the probability required, namely that of throwing the point 16 with 4 dice, is $\frac{125}{1296}$.

22. In the numerical solution of questions of this sort, it sometimes happens that the labour may be abridged by computing the probability of throwing a different point from that which is proposed, but which has the same number of chances in its favour. For example, let it be proposed to determine the probability that in throwing 10 dice the sum of the points will be 50. In this case, the smallest number of points that can possibly be thrown is 10, and the greatest 60; and the chances in favour of throwing 10 and of throwing 60 are obviously equal. The probability of throwing any given number of points above 10 is also evidently the same as that of throwing the number which is as much under 60; and consequently the probability of throwing 50 is the same as the probability of throwing 20, these numbers being at equal distances from the extremes. Now to find the probability of throwing 20 with 10 dice, or, which is the same, the probability that in 10 successive drawings from an urn

containing 6 balls, marked with the numbers 0, 1, 2, 3, 4, 5, the sum of the numbers drawn will be 10, we have $h=10$, $n=5$, $s=10$; whence $f=19$, $f'=13$, $f''=7$, f''' negative, and $k=1 \cdot 2 \dots 9$. Substituting these numbers in the series for N, and observing that since $f''-h+2=-1$, the third term becomes negative, we have

$$N = \frac{19 \cdot 18 \cdot 17 \cdot 16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9} (=92378) \\ - \frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9} (= -7150),$$

and consequently $N=85228$. Dividing this by $(n+1)^h=6^{10}=60466176$, the probability of throwing 20, or of throwing

50, with 10 dice is found $= \frac{85228}{60466176}$, or between $\frac{1}{709}$ and

$$\frac{1}{710}.$$

23. The probability that the whole number of points drawn in h trials will *not exceed* s is found by substituting for s the different values 0, 1, 2..... s in the series for N, and taking the sum of the results. This labour, however, may be avoided by means of a property of the figurate numbers. It is well known that the sum of the series of numbers obtained by giving n every value from $n=1$ to $n=v$ (v being any number whatever) in the formula

$$\frac{n(n+1)(n+2).....(n+u)}{1 \cdot 2 \cdot 3 \cdot u+1}$$

is expressed by this other formula

$$\frac{v(v+1)(v+2).....(v+u)(v+u+1)}{1 \cdot 2 \cdot 3 \cdot(u+1)(u+2)};$$

or, which is the same thing, that the sum of the series obtained by giving x successively every value from $x=1+u$ to $x=1+u+v$ in the formula

$$\frac{x(x-1)(x-2)\dots(x-u)}{1.2.3\dots u+1}$$

is expressed by this other formula

$$\frac{(x+1)x(x-1)\dots x-u}{1.2.3\dots u+2'}$$

in which $x=1+u+v$.

Comparing the different terms of the series for $N(20)$ with these last formulæ, it will be evident, that on giving s every value successively, from 0 to s in the value of f , and denoting by N' the sum of all the results, we shall have

$$\begin{aligned} N' = & + \frac{(f+1)f(f-1)(f-2)\dots(f-h+2)}{1.2.3.4\dots h} \\ & - \frac{(f'+1)f'(f'-1)(f'-2)\dots(f'-h+2)}{1.2.3.4\dots h} \times h \\ & + \frac{(f''+1)f''(f''-1)(f''-2)\dots(f''-h+2)}{1.2.3.4\dots h} \times \frac{h(h-1)}{1.2} \\ & - \&c. \end{aligned}$$

for the probability that the number of points thrown will not be greater than s .

As an example, let $h=10$, $s=5$, $n=5$; we have then $f=s+h-1=14$, $f'=f-(n+1)=8$, whence $f'-h+2=0$, and consequently the second term vanishes. Hence

$$N' = \frac{15.14.13.12.11.10.9.8.7.6}{1.2.3.4.5.6.7.8.9.10} = 3003.$$

The probability, therefore, of the sum of the numbers drawn in 10 drawings not exceeding 5, is $3003 \div (6)^{10}$, or

$\frac{3003}{60466176}$. And this is also the probability that in throwing 10 common dice the sum of the points does not exceed 16. (Simpson, p. 60.)

24. In the preceding questions the number of trials, de-

noted by h , has been supposed to be given; and the object, in every case, has been to determine the value of a given term, or of a given number of terms of the series produced by the development of the binomial $(p+q)^h$. But there is a numerous class of questions in respect of which the exponent h is unknown, and is required to be determined from the condition that an assigned term, or the sum of a certain number of assigned terms of the development, must have a given value. For example, let it be proposed to determine how often a common die must be thrown in order to give the probability of ace turning up *once at least*, equal to a given fraction u . Here the probability of throwing ace in any throw being $\frac{1}{6}$, we have $p=\frac{1}{6}$, and $q=\frac{5}{6}$. Now, as every term of the development of $(p+q)^h$, excepting the last, gives a combination in which ace occurs once or oftener, the question requires a value to be found for h , such, that the sum of the first h terms of that development, shall be equal to u . This may be done, in general, by the common methods of trial and error; but in the present case, the last term being the only one not included among those which contain a chance of throwing ace, it is evident that it is only necessary to find the last term alone in order to have the probability of *not* throwing ace in h trials, which by the question is $1-u$. The last term of the development is q^h ; therefore we must have the equation $q^h=1-u$; whence $h \log q = \log (1-u)$, and $h = \log (1-u) \div \log q$. Let $1-u = \frac{\beta}{\gamma}$ and $q = \frac{b}{c}$, we shall then have $\log (1-u) = \log \beta - \log \gamma$, and $\log q = \log b - \log c$; whence

$$h = \frac{\log \beta - \log \gamma}{\log b - \log c}.$$

Substituting in this general formula the particular numbers

given in the question, namely $b=5$, $c=6$; and supposing $u=\frac{1}{2}$, and consequently $\beta=1$, $\gamma=2$, we have $h=\frac{\log 2}{\log 6-\log 5}$; whence, by computing from the logarithmic tables, $h=3.8$. From this it follows, that in four trials the probability of throwing ace once at least, is greater than the probability of not throwing it at all.

If the question had been to determine in how many throws with two dice one may undertake, on an equality of chance, to throw aces at least once, we should have had $p=\frac{1}{36}$, $q=\frac{35}{36}$, and consequently $b=35$, and $c=36$. Substituting these numbers in the general formula, and observing, that in this case also $\beta=1$, $\gamma=2$, we get $h=\frac{\log 2}{\log 36-\log 35}=24.6$. The probability of not throwing aces once is therefore greater than the opposite probability or that of throwing aces once or oftener, when the number of throws is 24, but less when the number is 25.

These two questions are celebrated in the early history of the theory of Probability, from the circumstance that the Chevalier de Méré, by whom they were proposed to Pascal, declared the two results above stated to be inconsistent with each other, and thence took occasion to question the accuracy of the theory of combinations by means of which they had been obtained. He reasoned thus: Since the probability of throwing ace with one die is $\frac{1}{6}$, and that of throwing aces with two dice $\frac{1}{6}$ of $\frac{1}{6}=\frac{1}{36}$; therefore, if there be a given probability in favour of throwing ace in four throws with one die, there must likewise be the same probability of throwing aces with two dice in $6 \times 4=24$ throws; in other words, the chances in favour of an event E in a single trial, being six times more numerous than those in favour of F, there

will be as many chances in favour of F in six trials as there are in favour of E in one. The error consists in supposing that the number of trials must increase or diminish exactly in the inversed ratio of the probability of obtaining the proposed point.¹

25. The general question may be enunciated as follows : Let p = the probability an event E will happen, q the probability it will fail ; how many trials are required to give a probability = u that E will happen k times.

Let x = the number required. Taking the sum of all the terms of the development of $(p+q)^x$ in which the exponent of p is less than k , we shall have the probability that the event does *not* happen k times in x trials. This sum must consequently be made equal to $1-u$; therefore, beginning the last term, and writing the terms in the reverse order, we have the equation

$$q^x + x q^{x-1} p + \frac{x(x-1)}{1 \cdot 2} q^{x-2} p^2 \dots + \frac{x(x-1) \dots (x-k+2)}{1 \cdot 2 \dots (k-1)} q^{x-k+1} p^{k-1} = 1-u.$$

Let $p = e q$, and this equation becomes

$$q^x \left\{ 1 + x e + \frac{x(x-1)}{1 \cdot 2} e^2 \dots + \frac{x(x-1) \dots (x-k+2)}{1 \cdot 2 \dots (k-1)} e^{k-1} \right\} = 1-u$$

from which the value of x may be found by the ordinary methods of converging series.

If $p=q=\frac{1}{2}$, then $e=1$; and if we also suppose $u=\frac{1}{2}$, and consequently $1-u=\frac{1}{2}$, the equation will become

$$1 + x + \frac{x(x-1)}{1 \cdot 2} \dots + \frac{x(x-1) \dots x-k+2}{1 \cdot 2 \dots k-1} = \frac{1}{2} (1+1)^x.$$

¹ Pascal, *Œuvres*, tom. iv. p. 367 ; Lacroix, *Elémentaire*, p. 36.

But the first side of this equation is the expansion of $(1 \times 1)^x$ continued to k terms; therefore, since the sum of the first k terms is equal to one-half of the whole series, and the terms of the first half of the series are the same as those of the last, it follows that the whole number of terms must be $2k$. But the whole number of terms in the expansion of $(1 + 1)^x$ is $x + 1$; therefore $2k = x + 1$, and $x = 2k - 1$. Suppose $k = 10$, then $x = 19$; hence in tossing a shilling it is an even bet that head will turn up 10 times in 19 throws.

SECTION III.

OF THE PROBABILITY OF EVENTS DEPENDING ON A REPETITION OF TRIALS, OR COMPOUNDED OF ANY NUMBER OF SIMPLE EVENTS, THE CHANCES IN RESPECT OF WHICH ARE KNOWN *A PRIORI*, AND VARY IN THE DIFFERENT TRIALS.

26. Let us suppose the trials to consist in drawing balls from an urn containing a white balls, and b black balls, and that when a ball is extracted it is not returned to the urn. Make $a + b = c$, and let the extraction of a white ball be the event W , and that of a black ball the event B . At the first trial the probability of W is $\frac{a}{c}$ (4), and that of B , $\frac{b}{c}$. But at the second trial, the number of balls in the urn is diminished by 1; and the probability of drawing a white ball at the second trial is therefore not the same as it was in the first, but is influenced by the event which has already taken place. If W happened at the first trial, the number of white balls remaining in the urn is then $a - 1$; the number of black is b , and the number of both colours $c - 1$. The probability of W at the next trial is therefore $\frac{a-1}{c-1}$, and that of B is $\frac{b}{c-1}$. In like manner, if B happened at the first trial, the probability of W at the second is $\frac{a}{c-1}$,

and that of B is $\frac{b-1}{c-1}$. Hence (7) the different combinations

which can arise from two trials are the following :

WW, WB, BW, BB,

the probabilities of which are respectively,

$$\frac{a(a-1)}{c(c-1)}, \quad \frac{ab}{c(c-1)}, \quad \frac{ba}{c(c-1)}, \quad \frac{b(b-1)}{c(c-1)}.$$

Now if we neglect the order of succession in the two cases in which W and B are combined, the probability of the compound event which consists of the extraction of a ball

of each colour in the two trials, is $\frac{2ab}{c(c-1)}$, and the probabi-

lities of the three possible combinations are respectively :—

$$\frac{a(a-1)}{c(c-1)}, \quad \frac{2ab}{c(c-1)}, \quad \frac{b(b-1)}{c(c-1)}.$$

Comparing these with the probabilities of the same combinations when the chances are constant, or the ball is returned to the urn after each drawing, namely

$$\frac{a^2}{c^2}, \quad \frac{2ab}{c^2}, \quad \frac{b^2}{c^2},$$

the analogy of the two cases is obvious.

After two balls have been drawn, the whole number remaining in the urn is $c-2$; but the number of each colour depends on the two events that have already occurred. If two white balls have been drawn, the probability of drawing a white ball at the next trial will be $\frac{a-2}{c-2}$; but we have

just seen that the probability of WW is $\frac{a(a-1)}{c(c-1)}$; therefore

(7) the probability of WWW is $\frac{a(a-1)(a-2)}{c(c-1)(c-2)}$. The pro-

bability of drawing a black ball after two white have been drawn is $\frac{b}{c-2}$ (for there are now $c-2$ balls in the urn, of which b are white); therefore the probability of WWB is $\frac{a(a-1)b}{c(c-1)(c-2)}$. On forming in this manner all the different possible combinations which can result from three trials, we find

$$\begin{aligned}
 \text{WWW, probability,} &= \frac{a(a-1)(a-2)}{c(c-1)(c-2)}, \\
 \text{WWB, } &= \frac{a(a-1)b}{c(c-1)(c-2)}, \\
 \text{WBW, } &= \frac{ab(a-1)}{c(c-1)(c-2)}, \\
 \text{BWW, } &= \frac{ba(a-1)}{c(c-1)(c-2)}, \\
 \text{BBW, } &= \frac{b(b-1)a}{c(c-1)(c-2)}, \\
 \text{BWB, } &= \frac{ba(b-1)}{c(c-1)(c-2)}, \\
 \text{WBB, } &= \frac{ab(b-1)}{c(c-1)(c-2)}, \\
 \text{BBB, } &= \frac{b(b-1)(b-2)}{c(c-1)(c-2)}.
 \end{aligned}$$

If we disregard the order of succession in those combinations into which W and B both enter, and consider the occurrence of W twice and B once as the same compound event; and also the occurrence of W once and B twice as the same compound event, in whatever order they occur, the probability of the former will be $\frac{3a(a-1)b}{c(c-1)(c-2)}$, and of the latter $\frac{3ab(b-1)}{c(c-1)(c-2)}$.

27. In general, if $m' + n'$ balls have been drawn, of which m' have been found to be white and n' black, the number of white balls in the urn will now be $a - m'$, the number of black $b - n'$, and the whole number of both colours $c - m' - n'$. Hence the probability of drawing a white ball in the next

trial will be $\frac{a - m'}{c - m' - n'}$; and that of drawing a black ball

$\frac{b - n'}{c - m' - n'}$. Now if in these two fractions we substitute

successively for m' and n' all the different numbers from 0 to $m - 1$ and $n - 1$ respectively, the product of the $m + n$ numbers thus obtained will (7) be the probability of drawing m white balls and n black in an assigned order, in $m + n$ trials. Let this probability be denoted by K , and we shall have

$$K = \frac{a(a-1)(a-2)\dots(a-m+1) \times b(b-1)(b-2)\dots(b-n+1)}{c(c-1)(c-2)\dots(c-m-n+1)}$$

whatever the given order may be. Hence, if we denote by P the probability of m white balls and n black being drawn in any order whatever, in h trials, we shall have $P = UK$,

$$\text{where, as in (12), } U = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n}$$

the co-efficient of that term of the binomial $(p + q)^h$ which has for its argument $p^m q^n$; this co-efficient expressing all the different arrangements which can be formed of m things of one kind, and n things of another.

28. When the urn is supposed to contain balls of more than two different colours, the probability of any proposed number of each colour being drawn in a given number of trials is found with the same facility. Suppose it to contain a_1 of the first colour, a_2 of the second, a_3 of the third, and so on, and let $a_1 + a_2 + a_3 \text{ \&c.} = c$; then the probability that

in $m+n+r+\&c.=h$ trials, there will be drawn m of the first colour, n of the second, r of the third, $\&c.$ is

$$\begin{aligned} &U' \times a_1(a_1-1)(a_1-2)\dots\dots(a_1-m+1) \\ &\quad \times a_2(a_2-1)(a_2-2)\dots\dots(a_2-n+1) \\ &\quad \times a_3(a_3-1)(a_3-2)\dots\dots(a_3-r+1) \\ &\quad \times a_4(a_4-1)(a_4-2)\dots\dots(a_4-s+1) \\ &\quad \times \&c. \\ &\quad \div c(c-1)(c-2)\dots\dots(c-h+1) \end{aligned}$$

where, as in (15),

$$U' = \frac{1.2.3\dots\dots h}{1.2.3\dots\dots m \times 1.2.3\dots\dots n \times 1.2.3\dots\dots r \times \&c.}$$

29. The following examples will shew the use of the preceding formulæ.

Suppose a bag to contain 16 balls, of which 8 are white and 8 black, what is the probability that in drawing 8 balls from the bag the whole of them will be white ?

Applying the formula (27) to the solution of this question, we have $a=8$, $b=8$, $c=16$, $m=8$, $n=0$, and as the probability required is that of drawing white balls only, b cannot enter into any of the factors of the numerator; hence

$$K = \frac{8.7.6.5.4.3.2.1}{16.15.14.13.12.11.10.9} = \frac{1}{12870};$$

and since $m=h$, $U=1$, the probability sought, is therefore

$$\frac{1}{12870}.$$

Let there be a heap of 20 cards, wherein are 7 diamonds, 6 hearts, 4 spades, and 3 clubs; required the probability that in drawing 8 of them at a venture there shall come out 3 diamonds and 2 hearts? (Simpson, p. 21.)

The probability required in this case being that of drawing 3 diamonds, 2 hearts, and 3 other cards which are nei-

ther diamonds nor hearts, the spades and clubs may be considered as forming one parcel, containing 7 cards. We have then in the formula (28) $a_1=7$, $a_2=6$, $a_3=7$, $c=20$; $m=3$, $n=2$, $r=3$, $h=8$; therefore

$$U' = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3 \times 1 \cdot 2 \times 1 \cdot 2 \cdot 3} = 560,$$

and the probability required becomes,

$$560 \times \frac{7 \cdot 6 \cdot 5 \times 6 \cdot 5 \times 7 \cdot 6 \cdot 5}{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15 \cdot 14 \cdot 13} = \frac{1225}{3978}.$$

The odds against the event are therefore 2753 to 1225, or nearly 9 to 4.

Let 4 cards be drawn from a pack of 52; what is the probability of drawing one of each sort?

In this case we have $a_1=13$, $a_2=13$, $a_3=13$, $a_4=13$, $c=52$; also $m=1$, $n=1$, $r=1$, $s=1$, $h=4$, whence $U' = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 1 \cdot 1 \cdot 1} = 24$, and the probability required becomes, on

substituting these numbers in the formula (28),

$$24 \times \frac{13 \cdot 13 \cdot 13 \cdot 13}{52 \cdot 51 \cdot 50 \cdot 49} = \frac{2197}{20825} = \frac{1}{9} \text{ nearly.}$$

The odds against this event are nearly 8 to 1.

30. The following question, proposed by Huygens, and solved by Demoivre and Bernoulli (*Ars Conjectandi*, p. 59), belongs to the class of problems now under consideration.

An urn contains 12 balls, of which 4 are white and 8 black. Three gamesters A, B, and C agree that the first who, blindfold, shall draw a white ball shall be the winner of the stakes. They also agree that A shall draw first, B second, C third, A fourth, and so on; and the balls drawn are not replaced in the urn. It is proposed to find their respective probabilities of winning.

Here the play terminates as soon as a white ball is drawn, and it must therefore terminate with the 9th trial, if not sooner, inasmuch as, after 8 black balls have been drawn, the urn will contain only white balls, and the probability of drawing a white ball at the next trial will become certainty. The question will therefore be solved, if we determine the probabilities of the play ending with the 1st, 2d, 3d, 4th, &c. games respectively, and take the sum of the probabilities of its ending with the 1st, 4th, and 7th, for the probability of A's winning; the sum of the probabilities of its ending with the 2d, 5th, and 8th, for the probability of B's winning; and the sum of the probabilities of its ending with the 3d, 6th, and 9th, for the probability of C's winning.

For the sake of rendering the solution more general, let a be the number of white balls in the urn, b the number of black, and let $a+b=c$. The probability of drawing a white ball at the first trial, or of the play ending with the first trial, is then $\frac{a}{c}$.

The probability of the play ending with the second trial is compounded of the probability of a black ball being drawn at the first trial, and a white at the second; and the probability of both events (26) is $\frac{ba}{c(c-1)}$.

The probability of the play ending with the third trial is compounded of three separate probabilities, namely, that a black ball will be drawn at the first trial; that a black ball will be drawn at the second; that a white ball will be drawn at the third;—and the probability of the concurrence of these events (26) is $\frac{b(b-1)a}{c(c-1)(c-2)}$.

In general the probability of a black ball being drawn in

$x-1$ trials successively, and a white ball at the x th is $\frac{b(b-1)(b-2)\dots(b-x+2)a}{c(c-1)(c-2)\dots(c-x+1)}$. Substituting for a , b , and c in this formula the numbers proposed by Huygens, we obtain in respect of the 1st, 4th, and 7th trials, or the probability in favour of A,

$$\frac{4}{12} + \frac{8 \cdot 7 \cdot 6}{12 \cdot 11 \cdot 10} \cdot \frac{4}{9} + \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7} \cdot \frac{4}{6} = \frac{77}{165};$$

in respect of the 2d, 5th, and 8th trials, or the probability in favour of B,

$$\frac{8}{12} \cdot \frac{4}{11} + \frac{8 \cdot 7 \cdot 6 \cdot 5}{12 \cdot 11 \cdot 10 \cdot 9} \cdot \frac{4}{8} + \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6} \cdot \frac{4}{5} \\ = \frac{53}{165},$$

and in respect of the 3d, 6th, and 9th trials, or the probability in favour of C,

$$\frac{8 \cdot 7}{12 \cdot 11} \cdot \frac{4}{10} + \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8} \cdot \frac{4}{7} + \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5} \cdot \frac{4}{4} = \frac{35}{165}.$$

The chances in favour of A, B, and C, are therefore proportional to the numbers 77, 53, 35, respectively.

If the condition of the play had been that the ball was to be returned to the urn after each trial, the chances in favour of the three gamesters would have been easily found by the formula (12) to be respectively as the numbers 9, 6, and 4.

SECTION IV.

OF MATHEMATICAL AND MORAL EXPECTATION.

31. In the theory of probability, the term *expectation* is used to denote the product found by multiplying the value of a casual benefit into the probability of the event on which it is contingent taking place. But the value of a benefit may be estimated either with respect to its absolute amount, or to the amount of relative advantage it affords the individual who receives it. This consideration has led to a distinction between mathematical and moral expectation. When we place the circumstances of the individual entirely out of consideration, and have regard merely to the abstract or absolute value of the benefit, the product of its amount by the probability of obtaining it is the *mathematical expectation* of the individual; but when a relative value is assigned to the benefit, the product of this relative value by the probability of obtaining it is called the *moral expectation*, because it is estimated by certain moral considerations respecting the circumstances or fortune of the individual in whose favour the expectation exists, on the principle that a sum of money which may be relatively of very little importance to a man in possession of a large fortune may be of great importance to another who is less favourably circumstanced. We shall first consider the *mathematical expectation*.

32. Suppose A and B to engage in play ; let p be the probability of A's winning a game, q the probability of B's winning it, and s a sum of money staked on the issue of the game. By the definition, the mathematical expectation of A is ps , and that of B is qs . Now if we suppose these expectations to be purchased by A and B, the sums they ought respectively to pay for them, or in other words to stake on the issue of the game, must be proportional to their respective expectations, in order that they may play on equal terms. Let therefore a be the sum staked by A, and b the sum staked by B, we have then $ps : qs :: a : b$, and consequently $pb = qa$. Now suppose $a + b = s$, or that the sum played for is the amount of the stakes ; then, since b is the sum A expects to gain, and p is the probability of his gaining it, pb is the mathematical value of A's expectation of gain. In like manner qa is the mathematical value of B's expectation of gain. Hence it follows, that when the sum staked by each is proportional to his probability of winning, the mathematical expectations of the two players are equal ; so that after the stakes have been placed, and before the event is decided, they might exchange places without advantage or disadvantage to either. It follows likewise, that since the sum which the one must gain is just that which the other must lose, the product qa , which is B's expectation of gain, may be regarded as A's expectation of loss ; or (if taken with a negative sign) as part of A's whole expectation, which then becomes $pb - qa$. But $pb - qa = 0$; whence the condition of A before the event is decided is not altered by the circumstance of his having staked on the issue of the play.

33. This conclusion at first sight appears paradoxical ; for it is certain, that after the stakes are placed, A must either gain the sum b or lose a , and therefore his fortune will of

necessity either be increased by the gain of his adversary's stake, or diminished by the loss of his own. The explanation depends on theorems which will afterwards be demonstrated relative to the repetition of trials, from which it results, that though in a single trial the player must either lose or gain, yet on multiplying sufficiently the number of games, a probability will at length be obtained, approaching as nearly to certainty as we please, that the sum gained or lost in the long run will not exceed a certain given fraction (which may be as small as we please) of the whole sum staked, provided the play is undertaken on terms of mathematical equality. But this indefinite repetition of the hazard is practically impossible; and innumerable cases may easily be imagined, in which an individual will be guided by other considerations than the mere mathematical value of the expectation in undertaking or declining a risk. A person of moderate fortune would scarcely be persuaded to risk L.500 for the expectation of gaining L.5, though the chances might be 100 to 1 in favour of the event which would produce that sum; but numbers would be found willing enough to pay L.5 for the expectation of gaining L.500, the chances being 100 to 1 against them. In both cases, however, the expectation would be purchased at its real abstract value. According to the formula of mathematical expectation, the man whose sole fortune consists of a lottery ticket which has an equal chance of turning up a prize of L.20,000 or a blank, is in an equally advantageous position as he who is in possession of L.10,000; yet no man of ordinary prudence, if offered his choice of the two states, would hesitate as to which he ought to give the preference. Common sense will prevent a man from risking a sum, the loss of which would be attended with great privations, even

when, mathematically speaking, the chances are considerably in his favour. It is also obvious that two individuals whose fortunes are very unequal cannot engage in play with the same advantage, although the chances in favour of each, in respect of a single game, are precisely the same. The one who has a large fortune can repeat the hazard so often as to obtain a probability almost equal to certainty that his loss will not amount to any given sum; whereas the other, who cannot continue the play in case of loss, runs the risk of being ruined. It is thus evident, that in a multitude of cases the abstract theory of probability is not alone sufficient to give the value of an expectation, and that in dealing with contingent events, an individual must be guided to a certain extent by considerations of relative advantage.

34. Various hypotheses have been imagined for the purpose of reducing such relative or moral considerations to accurate calculation; but that which appears the most natural, and applicable to the greatest number of cases, consists in supposing the relative value of any infinitely small sum to be directly proportional to its absolute value, and inversely as the fortune of the individual who has an expectation of receiving it. This principle was first proposed by Daniel Bernoulli in the *Petersburg Commentaries* (vol. v.), and is there applied by him to the solution of a number of questions of great practical interest.

Let x be the absolute value of the capital, or, as it is denominated by Laplace, the physical fortune, of an individual; then, according to the hypothesis of Bernoulli, the moral advantage which he derives from an infinitely small increment of fortune $= dx$, is measured by the expression $c \frac{dx}{x}$, c being a constant to be determined by the nature of the ques-

tion. Now, if we suppose the physical fortune to arise from the accumulation of the elements dx , and denote by y the relative or *moral value* of the fortune, of which the absolute or physical value is x , we shall have

$$y = \int c \frac{dx}{x} = c \log. x + \text{constant}.$$

To determine the constant, we may suppose $y=0$, when x has a given value $=a$; this gives $0=c \log. a + \text{constant}$, whence $y=c (\log. x - \log. a)$, or $y=c \log. \frac{x}{a}$; and it is to be observed, that those values of x and y can never become negative, for as Bernoulli has remarked, it is only the person who is dying of hunger that can be said to possess absolutely nothing. In every other circumstance the mere possession of existence may be accounted a moral advantage, to which, however, it would be absurd to attempt to assign a numerical value.

35. From the above formula, it is easy to deduce a numerical expression for the value of a moral expectation. Let a be the original fortune of the individual, and α, β, γ , &c. sums to be received on the occurrence of certain contingent events, E, F, G, &c. This being supposed, if the event E happens, the absolute fortune of the individual becomes $a+\alpha$, and its relative value, therefore, according to the formula, is $c \log. \frac{a+\alpha}{a}$. If F happens, his absolute fortune becomes $a+\beta$, to which the corresponding relative value is $c \log. \frac{a+\beta}{a}$; and so on. Now, let the probabilities of the events E, F, G, &c. be respectively p, q, r , &c. (assuming $p+q+r+\&c.=1$, so that one or other of the events will necessarily happen), and let Y represent the relative fortune of

the individual arising from his expectation, then, since the value of a benefit in expectation is equal to the amount of the benefit multiplied by the probability of obtaining it, we have

$$Y=c \left\{ p \log. \frac{a+a}{a} + q \log. \frac{a+\beta}{a} + r \log. \frac{a+\gamma}{a} + \&c. \right\}$$

Let also X denote the absolute value of Y; then, by the formula, we have $Y=c \log. \frac{X}{a}$. On comparing these two values of Y, we get

$$\log. \frac{X}{a} = p \log. \frac{a+a}{a} + q \log. \frac{a+\beta}{a} + r \log. \frac{a+\gamma}{a} + \&c.;$$

and on passing to numbers,

$$\frac{X}{a} = \frac{(a+a)^p (a+\beta)^q (a+\gamma)^r, \&c.}{a^{p+q+r+\&c.}},$$

therefore, since $p+q+r+\&c.=1$,

$$X=(a+a)^p (a+\beta)^q (a+\gamma)^r, \&c.$$

In this expression X denotes the absolute value of the original fortune and of the expectation added together; if, therefore, we deduct a from X, the difference will be the value of the expectation, or the sum which, if it were to be received certainly, would procure the individual the same relative advantage as his expectation.

36. If the sums $a, \beta, \gamma, \&c.$ are supposed to be very small in comparison of a , so that quantities of the order $\left(\frac{a}{a}\right)^2$ may be neglected, the preceding equation becomes

$$X=a^{p+q+r+\&c.} + a^{p+q+r+\&c.-1} \left\{ pa + q\beta + r\gamma + \&c. \right\}$$

whence, since $p+q+r+\&c.=1$,

$$X=a+pa+q\beta+r\gamma+\&c.$$

Deducting from this the original fortune a , the remainder $pa + q\beta + r\gamma + \&c.$ is the value of the expectation, or the sum equivalent to the moral advantage. But the value of the mathematical expectation of the benefits $a, \beta, \gamma, \&c.$ of which the probabilities are respectively $p, q, r, \&c.$ is also $pa + q\beta + r\gamma + \&c.$ (31), therefore, when the contingent benefits are very small in comparison of the original fortune, the moral advantage and the mathematical expectation are sensibly the same.

37. From the formula $X = (a + a)^p (a + \beta)^q (a + \gamma)^r \&c.$ Bernoulli deduces the consequence that gambling or betting is attended with a moral disadvantage, even when the chances of gain or loss, mathematically speaking, are perfectly equal. To shew this, he proposes the following question. A, whose fortune is 100 crowns, bets 50 crowns with B, on the issue of an event of which the probability is $\frac{1}{2}$, on these terms: if the event happens, A is to receive from B 50 crowns; if it fails, he is to pay B 50 crowns; what is the relative value of A's fortune, after undertaking the bet, and before the event is decided? In this case, we have $a = 100, a = 50, \beta = -50, \gamma = 0$; also $p = \frac{1}{2}, q = \frac{1}{2}, r = 0$; and the formula (35) becomes

$$X = (100 + 50)^{\frac{1}{2}} \times (100 - 50)^{\frac{1}{2}},$$

whence $X = \sqrt{150 \times 50} = 87$; and, consequently, the condition of A is worse by 13 crowns than it was before he hazarded the bet. The moral disadvantage is therefore equivalent to this sum, though the terms of the play, according to the mathematical theory, are equal.

38. The conclusion arrived at in this particular case is easily shewn to be universally true. Let a be the capital of the player, p his probability of winning, q his probability

of losing, and s the sum at stake. In order that he may play on terms of mathematical equality, the part of the stakes contributed by himself, or the sum which he can lose, must be ps (32), and the part contributed by his adversary, or that which he may gain, must be qs . The equation in (35) therefore becomes

$$X = (a + qs)^p \times (a - ps)^q,$$

and if it can be shewn that this value of X is less than a , it will follow that his condition is rendered worse in consequence of having staked on the game. Now, dividing by a , and taking the logarithm of both sides of the equation, we get $\log. \frac{X}{a} = p \log. \left(1 + \frac{qs}{a}\right) + q \log. \left(1 - \frac{ps}{a}\right)$, the differential of which (making s variable) is

$$d \log. \frac{X}{a} = pqds \left(\frac{1}{1 + \frac{qs}{a}} - \frac{1}{1 - \frac{ps}{a}} \right).$$

But the second side of this equation is evidently negative; therefore $d \log. X \div a$ is negative; consequently the logarithm of $X \div a$ is negative, and X must be less than a . In all cases, therefore, the bet, if on even terms, produces a moral disadvantage.

39. Another consequence deduced by Bernoulli from this theory of moral expectation, is, that when property of any kind is exposed to a risk or hazard, it is more advantageous to expose it in parts to several risks independent of each other, than to expose the whole at once to a single risk, although the probability of loss be in both cases precisely the same. To prove this, he takes the following example. A merchant has a capital of L.4000, besides goods of the value of L.8000, which must be transported by sea. The probability of the loss of a vessel in the voyage being $\frac{1}{10}$, let it be

proposed to find the value of the moral expectation of the merchant in the case of the goods being embarked in a single vessel, and also in the case of one half being embarked in one vessel and the other half in another. Supposing the merchandise embarked in one ship, the absolute fortune of the merchant will be increased to L.12,000 in the event of the safe arrival of the ship, and will be reduced to L.4000 in the event of its being lost. The probability of the first of these events is $\frac{9}{10}$, and of the second $\frac{1}{10}$; therefore his absolute fortune becomes, in virtue of his expectation,

$$X = (12,000)^{\frac{9}{10}} \times (4000)^{\frac{1}{10}},$$

whence $X = 10751$. Deducting his other capital, L.4000, there remains L.6751 for the value of the moral expectation in respect of the venture.

Let us next suppose the merchandise embarked in equal parts in two ships. In this case there are three compound events to be considered, 1st, Both vessels may arrive in safety;

the probability of which is $\frac{9}{10} \times \frac{9}{10} = \frac{81}{100}$. 2d, One may arrive in safety and the other be lost; the probability of which, as it may happen in two ways, (11) is $2 \times \frac{9}{10} \times \frac{1}{10} = \frac{18}{100}$. 3d, Both may be lost; the probability of which is

$\frac{1}{10} \times \frac{1}{10} = \frac{1}{100}$. If the first of these events happen, the capital of the merchant will become $L.4000 + L.8000 = L.12,000$; if the second happen it will be $L.4000 + L.4000 = L.8000$; and if the third happen it will be only L.4000. With these numbers the formula becomes

$$X = (12,000)^{\frac{81}{100}} \times (8000)^{\frac{18}{100}} \times (4000)^{\frac{1}{100}},$$

whence $X=11033$. Deducting his other capital, which was exposed to no risk, there remains $L.7033$ for the value of the moral expectation. This sum exceeds the former by $L.282$; and it is easily found by following the same process of reasoning, that in proportion as the risk is divided among a greater number of ships, the moral expectation is increased, and approaches its limit, which is the value of the mathematical expectation, or $\frac{1}{3}$ of $L.8000=1.7200$.

40. The theory of moral expectation enables us likewise to assign the circumstances in which it is advantageous or otherwise, to insure property against particular hazards. There are three principal questions to be considered in reference to this subject; 1. The amount of premium the insured may pay without disadvantage; 2. The ratio of his fortune to the value of the sum exposed to risk, in order that it may be advantageous to insure at a given premium; and 3. The capital which the insurer or underwriter ought to possess, in order that he may insure a given risk with probable advantage to himself, and safety to the insured.

Let s be the value of a cargo which a merchant embarks in a ship, p the probability of the safe arrival of the vessel, and a his capital independently of s . The mathematical value of the premium for insurance is qs ; for, if we denote the premium by y , then y is the sum the insurer will gain if the vessel reaches its destination in safety, and $s-y$ is the sum he will lose if it does not; and by the theorem for the mathematical expectation $py=q(s-y)$; whence, since $p+q=1$, $y=qs$. If, therefore, the merchant insures the cargo, his absolute fortune becomes $a+s-q s=a+ps$; and if he does not insure, it is the value of X in the equation $X=(a+s)^r a^t$. Hence it will be advantageous or otherwise to insure according as $a+ps$ is greater or less than $(a+s)^r a^t$. Now the logarithm

of the first of these expressions, or $\log(a+ps)$, is equivalent to the integral $\int \frac{pds}{a+ps}$; and the logarithm of the second, or $p \log(a+s) + q \log a$, is equivalent to $\int \frac{pds}{a+s}$; but since p is a proper fraction, $a+ps$ is less than $a+s$, and therefore the first integral is greater than the second. Consequently $a+ps$ is, in general, greater than $(a+s)^p a^q$, and the insurance is attended with advantage. Let us now assume $x = a+ps - (a+s)^p a^q$, and x will be the sum the merchant could afford to pay the insurer above the mathematical value of the risk without moral disadvantage. If he pays less than $qs+x$, his relative fortune is increased by insuring; and if he pays more he is a loser. In practice the premium may be considered as less than $qs+x$, but greater than qs ; so that while the insured pays more than the mathematical value of the risk, he gains a moral advantage by the transaction.

To solve the second question, let e be the premium demanded for insuring the amount s ; then, the other capital of the merchant being a , his fortune after being insured is $a+s-e$; while if he takes the risk on himself, its value becomes $(a+s)^p a^q$. If, therefore, the value of a be determined from the equation $a+s-e = (a+s)^p a^q$, we shall have the amount of capital he ought to possess in order that it may be *morally* a matter of indifference to him whether he insures or not. As an example, let the value of the merchandise, or s , be L.10,000, $e = \text{L.}800$, and $p = \frac{1}{20}$. The equation then becomes

$$a + 9200 = (a + 10,000)^{\frac{1}{20}} a^{\frac{19}{20}};$$

whence a is found by approximation $= 5043$. It follows,

therefore, that unless his other capital amount to L.5043, it would be disadvantageous to neglect insuring, although the premium demanded exceed the mathematical value of the risk (which is $\frac{1}{20} \times \text{L.}10,000 = \text{L.}500$) by L.300.

The third question, the amount of capital the underwriter ought to possess, is determined precisely in the same way. Let b be his capital. After accepting the risk of the sum s for the premium e , his capital will become $b+e$ in the case of the vessel arriving in safety, and $b-s+e$ in the case of its being lost. The formula of the moral expectation therefore becomes $X = (b+e)^p(b-s+e)^q$; and in order that there may be neither advantage nor disadvantage in undertaking the risk, this value of X must be equal to his original capital, b . Supposing, therefore, s, e, p, q , to have the same significations as above, the equation from which b is to be determined is $b = (b+800)^{\frac{1}{2}}(b-9200)^{\frac{1}{2}}$, whence $b = 14243$. Unless, therefore, the capital of the insurer amounts to L.14,243, there would be a moral disadvantage in undertaking the risk of insuring a cargo worth L.10,000 for a premium of L.800; and it is easy to see, that if a smaller premium were demanded, the capital ought to be still greater. On making $e=600$, (which still exceeds the mathematical value of the risk), the value of b becomes L.29,878. Hence it follows, that a company possessing a large capital may not only with safety engage in speculations which might prove ruinous to another whose resources are more limited, but even derive from them a sure profit.¹

41. The theory of moral expectation which we have now been considering had its origin in a problem proposed by

¹ See the *Commentarii Acad. Petropolitanae*, tom. v.; Laplace, *Théorie des Prob.* p. 432; Lacroix, *Traité Élémentaire*, p. 132.

Nicolas Bernoulli to Montmort, which, from its having been discussed at great length by Daniel Bernoulli in the *Petersburg Memoirs*, has been usually called the Petersburg problem. It is this; A and B play at heads and tails. A agrees to pay B 2 crowns if head turn up at the first throw; 4 crowns if it turn up at the second, and not before; 8 if it turn up at the third, and not before; and, in general, 2^n crowns if it turn up at the n th throw, and not before: required the value of B's expectation? Here the probability of head turning up at the first throw is $\frac{1}{2}$; the probability of its turning up at the second, and not at the first, is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$; the probability of its not turning up either at the first or second, and of its turning up at the third, $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$, and so on. Hence the probabilities of B receiving 2, 4, 8, 16..... 2^n crowns

are respectively $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots, \frac{1}{2^n}$, consequently (31)

the mathematical value of B's expectation is

$$\frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \frac{1}{16} \times 16 \dots + \frac{1}{2^n} \times 2^n \text{ crowns.}$$

Now, as no limit can be assigned to n , inasmuch as it is possible that head may not turn up till after a very great, or any assignable number, of throws, this series, of which each term is unity, may go on for ever, and consequently the value of B's expectation becomes infinite. Yet it is obvious that no one would pay any considerable sum for the expectation. This disagreement between the dictates of common sense and the results of the mathematical theory, appeared to Montmort to involve a great paradox; although the question differs in this respect from no other question of chances in which the contingent benefit is very great, and

the probability of receiving it very small. If the play could be repeated an infinite number of times, B might undertake to pay without disadvantage any sum, however large, for his expectation. A result, however, more in accordance with ordinary notions, is obtained from the principle of Bernoulli. Let a be the amount of B's fortune before the play begins, x the value of his expectation, or the sum he pays A in consideration of the agreement, and make $z = a - x$. If head turn up at the first throw, B's fortune becomes $z + 2$; if at the second, and not before, $z + 2^2$; if at the third, and not before, $z + 2^3$; and so on. But the probabilities of these events being respectively $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots, \frac{1}{2^n}$, the formula for the moral expectation becomes (35)

$$X = (z + 2)^{\frac{1}{2}}(z + 2^2)^{\frac{1}{4}}(z + 2^3)^{\frac{1}{8}} \dots (z + 2^n)^{\frac{1}{2^n}}.$$

Now the sum which B ought to pay will be determined by making the value of his moral expectation, after the bet, and before the play begins, equal to his previous fortune; we have therefore $a = X$, that is,

$$a = (z + 2)^{\frac{1}{2}}(z + 2^2)^{\frac{1}{4}}(z + 2^3)^{\frac{1}{8}} \dots$$

The general term of this series being $(z + 2^n)^{\frac{1}{2^n}} = 2^{\frac{n}{2^n}} \left(1 + \frac{z}{2^n}\right)^{\frac{1}{2^n}}$, the equation may be put under the form

$$a = (2^{\frac{1}{2}} + 2^{\frac{2}{4}} + 2^{\frac{3}{8}} + 2^{\frac{4}{16}} \dots) \times \left(1 + \frac{z}{2}\right)^{\frac{1}{2}} \left(1 + \frac{z}{4}\right)^{\frac{1}{4}} \left(1 + \frac{z}{8}\right)^{\frac{1}{8}} \dots;$$

and since the logarithm of the first factor of this expression

$$\text{is } \left(\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{4}{16} + \&c.\right) \log 2, \text{ or } = \frac{1}{2} \left\{ 1 + 2\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + 4\left(\frac{1}{2}\right)^3 + \&c. \right\} \log 2 = \frac{1}{2} \left(1 - \frac{1}{2}\right)^{-2} \log 2 =$$

$$2 \log 2, \text{ we have } \log a = 2 \log 2 + \frac{1}{2} \log \left(1 + \frac{z}{2}\right) + \frac{1}{4} \log \left(1 + \frac{z}{4}\right) + \frac{1}{8} \log \left(1 + \frac{z}{8}\right) + \&c.$$

from which a value of z may be found by trial and error for any given value of a . Suppose $z=100$; on computing the first 10 terms of the series there results $a=107.89$, whence (since $x=a-z$) $x=7.89$; that is to say, if B possessed only 100 crowns before beginning the play, it would be morally disadvantageous for him to risk 8 crowns for the expectation, although its mathematical value be infinitely great. If we suppose $z=1000$, the sum of 11 terms gives $a=1011$, nearly; so that if B possessed a fortune of 1011 crowns, the value of the moral expectation would, to him, be about 11 crowns.

It is scarcely necessary to remark, that the results deduced from the principle of Bernoulli are of a character widely different from those which are calculated according to the mathematical expectation. The latter gives the precise value of a contingent benefit, without any assumption or hypothesis respecting the personal circumstances of the individual who may gain or lose it; whereas the considerations of relative advantage, of which it is the object of Bernoulli's theory to take account, are entirely arbitrary, and by their very nature incapable of being made the subject of accurate computation. It is evidently impossible to have regard to, or appreciate, all the circumstances which may render the same sum of money a more important benefit to one man than to another; and consequently every rule that can be given for the purpose must be liable to numerous exceptions. The principle, however, is thus far valuable, that it gives in the most common cases a plausible and judicious estimate of

the value of things which are not susceptible of exact appreciation ; and it has the advantage of being readily submitted to analysis. A different principle, proposed by the celebrated naturalist Buffon, consists in making the value itself of a casual benefit, instead of its infinitely small elements, inversely proportional to the fortune of the expectant ; but as this hypothesis has seldom been adopted, it is unnecessary to discuss it in this place.

SECTION V.

OF THE PROBABILITY OF FUTURE EVENTS DEDUCED FROM
EXPERIENCE.

42. In the preceding part of this article it has been assumed, in every case, that the number of chances favourable and unfavourable to the occurrence of a contingent event is known *a priori*, and consequently, that the probability of the event, or the ratio of the number of favourable cases to the whole number of cases possible, can be absolutely determined. But in numerous applications of the theory of probabilities, and these, generally speaking, by far the most important, the ratio of the chances in favour of an event to those which oppose it is altogether unknown; and we can form no idea of the probability of the event excepting from a comparison of the number of instances in which it has been observed to happen, with the whole number of instances in which it has been observed to happen and fail. In order to assign the probability of a contingent event in such cases, it is necessary to consider all the different causes or combinations of circumstances by which the event could possibly be produced, and to determine its probabilities successively on the hypotheses that each of these causes exists to the exclusion of all the others. The comparative facilities which these hypotheses give to the occurrence of the event which has actually arrived, will then enable us to de-

termine the relative probabilities of the different hypotheses, and consequently their absolute probabilities, since their sum is necessarily equal to unity; and when the probabilities of the different hypotheses, and of the occurrence of the event on each hypothesis, have been determined, the probability of the event occurring in a future trial will be found by the methods already explained.

43. Taking a simple case, let us suppose an urn to contain 4 counters, which are either white or black; that the number of each colour is unknown, but in four successive drawings (the counter drawn being replaced in the urn after each trial) a white counter has been drawn three times, and a black one once; and let it be proposed to assign the probability of drawing a counter of either colour at the next trial.

In the present case three hypotheses may be formed relative to the number of white and black counters in the urn. 1st, The urn may contain 3 white counters and 1 black; 2d, It may contain 2 white and 2 black; 3d, It may contain 1 white and 3 black; for a counter of each colour having been drawn, the other two possible cases, namely, that they are all white or all black, are excluded by the observation. Now, let p_1, p_2, p_3 , be the probabilities respectively of drawing a white counter on each hypothesis, and q_1, q_2, q_3 , the probabilities of drawing a black. Supposing the first hypothesis to be true, or that the compound event which has been observed was produced by the cause indicated by that hypothesis, we have $p_1 = \frac{3}{4}$, $q_1 = \frac{1}{4}$; and the probability of the observed event, or that 3 white counters and 1 black would be drawn, (12) is $4p_1^3 q_1 = \frac{27}{64}$. The second hypothesis gives $p_2 = \frac{1}{2}$, $q_2 = \frac{1}{2}$, whence $4p_2^3 q_2 = \frac{1}{8}$. The third hypothesis gives $p_3 = \frac{1}{4}$, $q_3 = \frac{3}{4}$, whence $4p_3^3 q_3 = \frac{27}{64}$. The probabilities

of the observed compound event, on each of the three hypotheses, are therefore, respectively, $\frac{2}{6}\frac{7}{4}$, $\frac{1}{6}\frac{6}{4}$, $\frac{5}{6}\frac{3}{4}$; and the question now arises, how are the probabilities of the different hypotheses to be estimated? As we have no data, *a priori*, for determining this question, we must assume the probabilities of the different hypotheses to be respectively proportional to the probabilities they severally give of the observed compound event; in other words, we must assume the probability of any hypothesis to be greater or less according as it affords a greater or smaller number of combinations favourable to the event which has been observed to take place. Thus, if C and C₁ be two independent causes from which an observed event E may be supposed to arise, and C furnishes 20 different combinations out of a given number, favourable to the occurrence of E, while C₁ furnishes only 10 such combinations out of the same number, we naturally infer that the probability of the cause C having operated to produce E, is twice as great as the probability that the event was produced by the operation of the cause C₁. Applying this principle to the present example, the probabilities of the three hypotheses are respectively proportional to the three fractions $\frac{2}{6}\frac{7}{4}$, $\frac{1}{6}\frac{6}{4}$, $\frac{5}{6}\frac{3}{4}$, or to the numbers 27, 16, 3; and as no other hypotheses are admissible, the sum of their probabilities must be unity; therefore, making π_1 the probability of the first hypothesis, π_2 that of the second, and π_3 that of the third, we have

$$\pi_1 = \frac{27}{46}, \quad \pi_2 = \frac{16}{46}, \quad \pi_3 = \frac{3}{46}.$$

44. Having found the probabilities of the different hypotheses, that of drawing a white counter at the next trial is obtained without difficulty; for according to what was

shewn in (9), the probability of this simple event must be equal to the sum of its probabilities relative to the different hypotheses, each multiplied into the probability of the hypothesis itself. Now it has been seen that, on the first hypothesis, the probability of drawing a white ball is $\frac{3}{4}$; on the second $\frac{2}{4}$, and on the third $\frac{1}{4}$; and that the probabilities of the hypotheses are respectively $\frac{27}{46}$, $\frac{16}{46}$, $\frac{5}{46}$; therefore the probability of a white counter being drawn at the next trial is

$$\frac{3}{4} \times \frac{27}{46} + \frac{2}{4} \times \frac{16}{46} + \frac{1}{4} \times \frac{5}{46} = \frac{116}{184}.$$

In like manner, the probability of a black counter being drawn at the next trial is

$$\frac{1}{4} \times \frac{27}{46} + \frac{2}{4} \times \frac{16}{46} + \frac{3}{4} \times \frac{5}{46} = \frac{68}{184};$$

and the sum of these two fractions is unity, as it ought to be, since the counter drawn must necessarily be white or black.

45. The reasoning which has been employed in this particular case is of general application. Let E be an observed event, simple or compound, of which the particular cause is unknown, but which may be ascribed to any one of the n causes, $C_1, C_2, C_3, \dots, C_n$, which, before the event has happened, are all equally probable, and such that the operation of any one of them excludes that of the others, so that the event E is produced by one of them alone, and not by the joint agency of several of them. Let the probabilities of the observed event E on the hypothesis that it has proceeded from each of those causes be respectively $P_1, P_2, P_3, \dots, P_n$, so that if the cause, for instance, C_i were the true one, the probability of the event E , previous to the observation, would be P_i ; and let the probabilities (as determined by the

event) of the existence of the different causes be respectively $\varpi_1, \varpi_2, \varpi_3, \dots, \varpi_n$. From the principle laid down in the preceding paragraph, namely, that the probabilities of the different causes or hypotheses are proportional to the probabilities they respectively give of the observed event, we have

$$\varpi_1 : \varpi_2 : \varpi_3 : \dots : \varpi_n :: P_1 : P_2 : P_3 : \dots : P_n$$

whence, making $P_1 + P_2 + P_3 + \dots + P_n = \Sigma P_i$, and observing that $\varpi_1 + \varpi_2 + \varpi_3 + \dots + \varpi_n = 1$ (since it is assumed that there are no other causes than those specified from which the event could arise), we have

$$\varpi_1 = \frac{P_1}{\Sigma P_i}, \varpi_2 = \frac{P_2}{\Sigma P_i}, \varpi_3 = \frac{P_3}{\Sigma P_i}, \dots, \varpi_n = \frac{P_n}{\Sigma P_i},$$

whence it appears that the probability of each hypothesis respecting the cause of the observed event is found by dividing the probability of the event on the supposition that that particular cause alone existed, by the sum of its probabilities in respect of all the causes. Let us now assume the probabilities of a future event E' (which may be the same with E or different, but depending on the same causes) in respect of the several hypotheses, to be, $p_1, p_2, p_3, \dots, p_n$; so that if the particular cause C_i be the true one, the probability of E' is p_i ; and let Π be the probability of E' in respect of all the causes, then by (9), Π will be equal to the sum of the probabilities $p_1, p_2, p_3, \dots, p_n$ relative to the different hypotheses, each multiplied by the probability of the hypothesis; that is to say we shall have

$$\Pi = p_1 \varpi_1 + p_2 \varpi_2 + p_3 \varpi_3 + \dots + p_n \varpi_n;$$

or $\Pi = \Sigma p_i \varpi_i$, the symbol Σ indicating the sum of all the different values of p and ϖ in respect of the different causes $C_1, C_2, C_3, \dots, C_n$.

46. It may be worth while to remark that the word cause is not here used in its ordinary acceptation to denote the combination of circumstances, physical or moral, of which the event is a *necessary* consequence. In the sense we have used the term, the cause C is that which gives rise to the determinate probability P , that the event E will happen; but so long as this probability falls short of certainty, its existence also implies that of another probability, $1-P$, that the contrary event F will happen. If we make $P=1$, the existence of the cause C would necessarily involve the occurrence of E ; and it is in this particular sense that the word cause is ordinarily used. In the theory of probabilities the causes of events are considered only in reference to the number of chances they afford for the occurrence of those events which they may possibly, but do not necessarily, produce.

47. The following example may serve to illustrate the method of applying the preceding formulæ. An urn contains n balls, which are known to be either white or black. A ball is drawn at random and found to be white; required the probability of drawing a white ball at the next trial?

In this case, the number of hypotheses that may be made respecting the contents of the urn, is n ; for we may suppose that it contained one, or two, or any number of white balls from 1 to n , and each of these cases may be considered as a distinct cause of the observed event E . Let these causes or hypotheses be $C_1, C_2, C_3, \dots, C_n$, and let us suppose the true cause was C_i , or that the urn contained i white balls. On this hypothesis the probability of the observed event E is $\frac{i}{n}$, whence $P_i = \frac{i}{n}$; and therefore, making i succes-

sively equal to 1, 2, 3, ..., n , we have $\Sigma P_i = \frac{1}{n}(1 + 2 + 3 + \dots + n)$

But the sum of this arithmetical series is $\frac{n(n+1)}{2}$, therefore

$\Sigma_i P = \frac{1}{2}(n+1)$, and consequently,

$$\varpi_i = \frac{P_i}{\Sigma P_i} = \frac{2i}{n(n+1)},$$

which is the probability of the assumption that the event proceeded from the cause C_i , or that the urn contained i

white balls. If we suppose $i=n$ we have $\varpi_n = \frac{2}{n+1}$ for the

probability that all the balls are white; and if we also suppose $n=3$, this becomes $\frac{1}{2}$; whence if an urn contain 3 balls which must be either black or white, and a white ball be drawn at the first trial, it is an even wager, after the trial, that all the balls are white.

48. Having found, from the observed event E , the probabilities of the different hypotheses, we have now to determine the probability Π of the event E' (the drawing of a white ball) at the next trial. Here two cases present themselves; according as the ball is replaced in the urn, or is not; or in general, according as the law of the chances remains constant during the series of trials or varies.

1st, Let us suppose that the ball has been replaced in the urn. In this case the probability of the event E' , on the hypothesis that the urn contains i white balls, is $\frac{i}{n}$; that is to

say $p_i = \frac{i}{n}$. But the probability ϖ_i of this hypothesis, as

found above, is $\frac{2i}{n(n+1)}$; therefore $p_i \varpi_i = \frac{2i^2}{n^2(n+1)}$; whence

the general formula (45) $\Pi = \sum p_i \pi_i$ becomes $\Pi = \sum \frac{2i^2}{n^2(n+1)} = \frac{2}{n^2(n+1)} \sum i^2$. Now $\sum i^2 = \sum i(i+1) - \sum i$. But by the property of the figurate numbers referred to in (23), the sum of the series of numbers obtained by giving i every value from $i=1$ to $i=n$ in the formula $\frac{n(n+1)}{1 \cdot 2}$ is expressed by $\frac{n(n+1)(n+2)}{1 \cdot 2 \cdot 3}$; therefore $\sum i(i+1) = \frac{n(n+1)(n+2)}{3}$.

We have also as above

$$\sum i = 1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2};$$

consequently,

$$\sum i^2 = \frac{n(n+1)(n+2)}{3} - \frac{n(n+1)}{2} = \frac{n(n+1)(2n+1)}{2 \cdot 3}$$

and therefore

$$\Pi = \frac{2}{n^2(n+1)} \times \frac{n(n+1)(2n+1)}{2 \cdot 3} = \frac{2n+1}{3n}.$$

2d, Suppose the ball which has been extracted is not replaced in the urn. In this case, on the hypothesis that the urn at first contained i white balls, the probability of drawing a white ball at the next trial is $\frac{i-1}{n-1}$; that is, $p_i = \frac{i-1}{n-1}$;

and the probability of the hypothesis is the same as in the former case, or $\pi_i = \frac{2i}{n(n+1)}$; therefore $p_i \pi_i = \frac{2i(i-1)}{(n-1)n(n+1)}$;

and consequently $\Pi = \sum p_i \pi_i = \frac{2}{(n-1)n(n+1)} \sum i(i-1)$. Now the value of $\sum i(i-1)$ will evidently be found by writing $n-1$ for n in the above expression for $\sum i(i+1)$; whence

$\Sigma i(i-1) = \frac{(n-1)n(n+1)}{3}$, and, therefore, in this case

$$\Pi = \frac{2}{(n-1)n(n+1)} \times \frac{(n-1)n(n+1)}{3} = \frac{2}{3}.$$

When n is a very large number, the ratio of $2n+1$ to $3n$, the value of Π in the former case, does not sensibly differ from $\frac{2}{3}$, and therefore in both cases $\Pi = \frac{2}{3}$. Hence it follows, that if an event, depending on unknown causes, can happen only in one of two ways, and it has been observed to happen once, the odds are two to one in favour of its happening in the same way at the next occurrence.

49. The expression for ϖ in (45) was determined on the supposition that previously to the experiments being made, we are entirely ignorant of the relative numbers of the two sorts of balls in the urn, and have no reason to suppose one hypothesis more probable than another. If, however, we happen to know, previously to the experiment, that the different causes $C_1, C_2, C_3, \&c.$ have not all the same number of chances in their favour, or that the probabilities of the different hypotheses have relative values, it becomes necessary to introduce those relative values, in consequence of which $\varpi_1, \varpi_2, \&c.$, will receive a modification. Let us conceive a number of urns, each containing balls of two colours, black and white, to be distributed in n groups, $A_1, A_2, A_3, \dots, A_n$, in such a manner that the ratio of the number of white balls to the number of black balls is the same in respect of each urn belonging to the same group, and consequently that the probability of drawing a ball of either colour is the same from whichever urn in the group it may happen to be drawn, but different in respect of the different groups; and let the probabilities of drawing a white ball from each of the different groups be re-

spectively $P_1, P_2, P_3, \dots, P_n$. Now, let us suppose there are a_1 urns in the group A_1 , a_2 in the group A_2 , and so on, and let $s =$ the whole number of urns, so that $s = a_1 + a_2$

$+ a_3 + \dots + a_n$; then, if we make $\frac{a_1}{s} = \lambda_1$, $\frac{a_2}{s} = \lambda_2$, and so

on, λ_1 will be the *a priori* probability that a ball drawn from any urn at random, will be drawn from the group A_1 ; λ_2 the probability it will be drawn from the group A_2 ; and, in general, λ_i the probability it will be drawn from the group A_i . This being premised, suppose a trial to be made, and that the event E is a white ball; the probability ϖ_i of the hypothesis that the ball was drawn from the group A_i is found as follows. The *a priori* probability of the ball being drawn from the group A_i is λ_i ; and if the ball is actually drawn from that group, the probability of its being white is P_i ; therefore the probability of both events is

$\lambda_i P_i$; and consequently (45), $\varpi_i = \frac{\lambda_i P_i}{\sum \lambda_i P_i}$, the symbol of summation Σ extending to all the values of i from $i=1$ to $i=n$.

50. In the applications of the theory to physical or moral events, the different groups of urns here imagined may be regarded as so many independent causes C_1, C_2, C_3 , &c. by any one of which the event E might have been produced; ϖ_i is the probability that the event was produced by the particular cause C_i ; P_i is the probability that the cause C_i if it had alone existed, would have produced the observed event E ; and λ_i is the probability, previously to the experiment, that C_i would be the efficient cause. The formula $\varpi_i = \frac{\lambda_i P_i}{\sum \lambda_i P_i}$, therefore, shews that the probability of any one of the possible causes (C_i) of an observed event is equal to the product of the probability (P_i) of the event taking place if

that cause acted alone multiplied into the probability λ_i that the cause C_i is the true one, and divided by the sum $(\sum \lambda_i P_i)$ of all the similar products formed relatively to each of the causes from which the event can be supposed to arise.

51. The formulæ now obtained can only be used when the number of hypotheses is finite ; but in the applications of the theory it most frequently happens that an infinite number of hypotheses may be made respecting the causes of an observed event, as would be the case in the above example if the number of balls in the urn had been unknown. In such cases, in order to find the values of ϖ and Π , it becomes necessary to transform the sums Σ into definite integrals, which is accomplished by means of the theorem $\Sigma X = \int_0^1 X dx$, where X is a function of x . Suppose a ball to have been drawn a great number of times in succession from an urn (the number in which is unknown) and replaced in the urn after each drawing, and that the result has been a white ball m times and a black ball n times, the probable constitution of the urn, and thence the probability of drawing a white ball at a future trial will be found as follows. Assume the hypothesis that the ratio of the number of white balls to the whole number in the urn is $x : 1$, and let ϖ be the probability of the hypothesis. On this hypothesis the probability of drawing a white ball in any trial is x , and that of drawing a black ball $1-x$, and consequently, the probability of drawing m white and n black in $m+n$ trials is $Ux^m(1-x)^n$ by (12). We have therefore for the probability of the observed compound event $P = Ux^m(1-x)^n$; whence in consequence of the above formula for transforming a sum into a definite integral $\Sigma P = U \int_0^1 x^m(1-x)^n dx$ (U being independent of x) and therefore

$$x = \frac{P}{\Sigma P} = \frac{x^m(1-x)^n}{\int_0^1 x^m(1-x)^n dx}.$$

The value of the integral in the denominator of this fraction is obtained by the usual method of integrating by parts.

Since

$$d \frac{x^{m+1}(1-x)^n}{m+1} = x^m(1-x)^n dx - \frac{n}{m+1} x^{m+1}(1-x)^{n-1} dx,$$

therefore

$$\int x^m(1-x)^n dx = \frac{x^{m+1}(1-x)^n}{m+1} + \frac{n}{m+1} \int x^{m+1}(1-x)^{n-1} dx,$$

In like manner we get $\int x^{m+1}(1-x)^{n-1} dx$

$$= \frac{x^{m+2}(1-x)^{n-1}}{m+2} + \frac{n-1}{m+2} \int x^{m+2}(1-x)^{n-2} dx.$$

Continuing this operation n times, or till the exponent of $(1-x)$ becomes $n-n=0$, the last integral will be

$$\int x^{m+n} dx = \frac{x^{m+n+1}}{m+n+1};$$

therefore, collecting the several terms into one sum, we have

$$\begin{aligned} \int x^m(1-x)^n dx &= \frac{x^{m+1}(1-x)^n}{m+1} + \frac{n x^{m+2}(1-x)^{n-1}}{(m+1)(m+2)} \dots\dots\dots \\ &\dots\dots\dots + \frac{n(n-1)(n-2)\dots\dots\dots 2 \cdot 1 \cdot x^{m+n+1}}{(m+1)(m+2)\dots\dots\dots m+n+1}. \end{aligned}$$

When $x=0$, all the terms of this series vanish, and when $x=1$ they all vanish excepting the last; therefore between the limits $x=0$ and $x=1$, the value of the integral is the last term of the series when x in that term $=1$; that is to say,

$$\int_0^1 x^m(1-x)^n dx = \frac{n(n-1)(n-2)\dots\dots 2 \cdot 1}{(m+1)(m+2)\dots\dots m+n+1}.$$

For the sake of brevity, let the symbol $[x]$ be adopted to

represent the continued product $1 \cdot 2 \cdot 3 \dots x$ of the natural numbers from 1 to x ,¹ whence by analogy $[x+y]$ will represent the continued product of the same series from 1 to the number denoted by $x+y$. Multiplying, then, the numerator and denominator of the above expression by $1 \cdot 2 \cdot 3 \dots m = [m]$, we get

$$\int_0^1 x^m (1-x)^n dx = \frac{[m][n]}{[m+n+1]};$$

whence the probability of the hypothesis, in consequence of the equation above found, becomes

$$\pi = \frac{[m+n+1]}{[m][n]} x^m (1-x)^n.$$

From this value of π we are enabled to deduce that of Π , the probability of drawing a white ball at the next trial. By (45) $\Pi = \Sigma \pi p$. Now, since by hypothesis the number of white balls in the urn is to the whole number of both colours in the ratio of x to 1, the probability of drawing a white ball is x ; consequently $p = x$, and therefore $\Pi = \Sigma \pi x = \int_0^1 \pi x dx = \frac{[m+n+1]}{[m][n]} \int_0^1 x^{m+1} (1-x)^n dx$. But the value of $\int_0^1 x^{m+1} (1-x)^n dx$ will evidently be obtained by substituting $m+1$ for m in the expression found for $\int_0^1 x^m (1-x)^n dx$. This substitution gives

$$\int_0^1 x^{m+1} (1-x)^n dx = \frac{[m+1][n]}{[m+n+2]},$$

whence, observing that $[m+1] \div [m] = m+1$, and $[m+n+2] \div [m+n+1] = m+n+2$, we have

$$\Pi = \frac{m+1}{m+n+2}.$$

¹ This convenient notation has been adopted by Mr. De Morgan.

The probability of the contrary event, or of drawing a black ball, is $1-\Pi = \frac{n+1}{m+n+2}$. As the numbers m and n become larger, these two fractions approach nearer and nearer to their limits $\frac{m}{m+n}$ and $\frac{n}{m+n}$, which are the *a priori* probabilities p and q of the respective events when the ratio of the number of white balls in the urn is to that of the black balls as m to n .

52. The probability of drawing m' white balls and n' black balls in $m' + n'$ future trials is found in a similar manner, and the problem may be thus stated. E and F are two contrary events, depending on constant but unknown causes; and it has been observed, that in $m+n=h$ successive instances the event E has occurred m times and F n times, required the probability that in $m' + n' = h'$ future instances, E will occur m' times and F n' times.

Assume, as in the last case, the facility of the occurrence of E to that of F to be in the ratio of x to $1-x$; we have then, as before, for the probability of the hypothesis, $\pi =$

$\frac{[h+1]}{[m][n]} x^m (1-x)^n$. Now on this hypothesis the probabi-

lity of E in the next instance is x , and that of F is $1-x$, whence the probability of m' times E and n' times F in the next h' trials being denoted by p , we have (12) $p = U' x^{m'}$

$(1-x)^{n'}$, making $U' = \frac{1 \cdot 2 \cdot 3 \dots h'}{1 \cdot 2 \cdot 3 \dots m' \times 1 \cdot 2 \cdot 3 \dots n'} =$

$\frac{[h']}{[m'][n']}$. We have therefore $\pi p = U' \frac{[h+1]}{[m][n]} x^{m+m'}$

$(1-x)^{n+n'}$ for the probability of the compound event on this hypothesis. To find its probability Π on the infinite

number of hypotheses formed by supposing x to increase by infinitely small increments from $x=0$ to $x=1$, we have

$\Pi = \Sigma \varpi p = \int_0^1 \varpi p dx$. On substituting for ϖp the value just

found, we get $\Pi = U' \frac{[h+1]}{[m][n]} \int_0^1 x^{m+m'} (1-x)^{n+n'} dx$, and

it is manifest that the value of this integral will be obtained by substituting $m+m'$ for m , and $n+n'$ for n in the value of $\int_0^1 x^m (1-x)^n dx$ found above. This substitution gives

$$\int_0^1 x^{m+m'} (1-x)^{n+n'} dx = \frac{[m-m][n+n']}{[h+h'+1]},$$

whence we conclude

$$\Pi = U' \frac{[m+m'] [n+n'] [h+1]}{[m][n][h+h'+1]}.$$

The most probable hypothesis will be found by making the value of ϖ a maximum, or its differential coefficient equal

to zero. Differentiating the equation $\varpi = \frac{[h+1]}{[m][n]} x^m (1-x)^n$,

and making $\frac{d\varpi}{dx} = 0$, we get $m(1-x) = nx$, whence $x =$

$\frac{m}{m+n}$. The most probable supposition, therefore, respect-

ing the contents of the urn is, that the two sorts of balls are in the same proportions as have been shewn by the previous drawings. We shall have further occasion for these formulæ when we come to consider the cases in which m and n are large numbers.

SECTION VI.

OF BENEFITS DEPENDING ON THE PROBABLE DURATION OF
HUMAN LIFE.

53. In applying the principles of the theory of probability to the determination of the values of benefits depending on life, the fundamental element which it is necessary to determine from observation is the probability that an individual at every given age within the observed limits of the duration of life, will live over a given portion of time, for instance one year; for when this has been determined for each year of age, the probability that an individual, or any number of individuals, will live over any assigned number of years, is easily deduced. Thus, if the probabilities that an individual A, whose age is y , will live over 1, 2, 3... x years, be denoted respectively by $p_1, p_2, p_3 \dots p_x$; and if $q_1, q_2, q_3 \dots q_x$ denote the same probabilities in respect of an individual whose age is $y+1$ years; $r_1, r_2, r_3 \dots r_x$, the same in respect of an individual whose age is $y+2$ years, and so on; then, since the probability p_2 which A has of living over 2 years is obviously compounded of the probability p_1 of his living over 1 year, and of the probability q_1 that, having attained the age $y+1$, he will live another year, we have, by (7), $p_2 = p_1 q_1$. Again, the probability p_3 that A will live over three years, being compounded of the proba-

bility p_2 that he will live over two years, and of the probability r_1 that, having attained the age $y+2$ years, he will survive another year, we have $p_3 = p_2 r_1 = p_1 q_1 r_1$. In like manner $p_4 = p_1 q_1 r_1 s_1$, and so on; so that the probabilities $p_2, p_3, p_4 \dots p_x$ are successively derived from p_1, q_1, r_1, s_1 , &c. which are supposed to be the data of observation.

If a large number n of individuals, all born in the same year, were selected, and if it were observed that the number of them remaining alive at the end of the first year is n_1 , at the end of the second year n_2 , at the end of the third n_3 , and so on, then the probabilities p_1, p_2, p_3 , &c. would be given directly by the observation, being respectively equal to the quotients $\frac{n_1}{n}, \frac{n_2}{n}, \frac{n_3}{n}$, &c. But the most

accurate observations of mortality are furnished by the experience of the annuity and assurance offices, where they are not made on an isolated number, diminishing, and consequently giving a less valuable result every year, but on a comparison of the numbers which, in a series of years, enter upon and survive each year of age. This observation gives p_1, q_1, r_1, s_1 , &c. whence p_2, p_3, p_4 , &c. are found, as above, for every year of life.¹

54. The values of annuities on lives, and of reversionary sums to be paid on the failure of lives, are found by combining the probabilities p_1, p_2, p_3 , &c. with the rate of interest of money. Let r = the rate of interest, that is to say, the interest of L.1 for a year, and v = the present value of L.1 to be received at the end of a year, we shall then have $v = 1 \div (1 + r)$. Now an annuity, payable yearly, is always understood in this sense, that the first payment becomes due

¹ For further details on this subject, see MORTALITY, vol. xv. p. 550.

at the end of a year after the annuity is created. Suppose then the annuity to be L.1, the present value of the first payment, if it were to be received certainly, is v ; but the receipt of this sum is contingent on the annuitant being alive at the end of the year, the probability of which we suppose to be p_1 ; therefore (7) the present value of L.1 subject to the contingency, is vp_1 . In like manner, the present value of L.1 to be received certainly at the end of x years is v^x ; but the annuity will only be received at the end of the x th year if the annuitant be then living, the probability of which is p_x ; therefore the present value of that particular payment is $v^x p_x$. Hence if A denote the present value of the annuity, or the sum in hand which is equivalent to all the future payments, we shall have $A = \Sigma v^x p_x$; the sum Σ including all values of x from $x=1$ to $x=$ the number for which $p=0$. If the annuity be a pounds, its value is obviously $= a \Sigma v^x p_x = aA$.

55. The series denoted by $\Sigma v^x p_x$ may be divided into two parts, $\Sigma v^n p_n + \Sigma v^x p_x$, where n is to be taken from 1 to n , and x from $n+1$ to the number for which p vanishes. The first gives the value of the *temporary* annuity on the given life for n years, and the second the value of the *deferred* annuity, that is to say, of the annuity to commence n years hence if the individual shall be then living, and to continue during the remainder of his life. Let A be the value of the annuity on the life of a person now aged y years for the whole of life, $A^{(tn)}$ the value of a temporary annuity on the same life for n years, and $A^{(dn)}$ the value of an annuity deferred n years on the same life, we have then $A = A^{(tn)} + A^{(dn)}$.

To find $A^{(dn)}$, let A_n be the value of an annuity on a life aged $y+n$ years. If the person now aged y years lives over n years, the value of an annuity on the remainder of his

life will *then* be A_n . The present value of this sum, if it were to be received certainly, is $v^n A_n$, and the probability of receiving it is p_n ; therefore its value is $v^n p_n A_n$. Hence

$$A^{(dn)} = v^n p_n A_n, \text{ and } A^{(tn)} = A - v^n p_n A_n;$$

so that the values of temporary and deferred annuities are readily computed from tables of A and p for all the different ages.

56. The equation $A = A^{(tn)} + A^{(dn)}$ gives a formula by which the values of A are readily deduced from one another. Let $n=1$; we have then $A = A^{(t1)} + v p_1 A_1$. But $A^{(t1)}$, the value of an annuity for one year, is merely the value of the first payment to be received in the event of the given life surviving one year. Its value is therefore $v p_1$; and we have consequently $A = v p_1 + v p_1 A_1$, or $A = v p_1 (1 + A_1)$. This formula, which gives the value of an annuity at any age in terms of the next higher age, and greatly facilitates the computation of the annuity tables, is due to Euler.

57. The value of an annuity on the joint lives of any number of individuals, that is, to continue only while they are *all* living, is calculated precisely in the same manner as the annuity on a single life. Let there be any number of individuals, A, B, C, D , &c. and let the probabilities of each living over one year be respectively p_1, q_1, r_1, s_1 , &c. and let P_1 be the probability that they will *all* live over one year; then

$$P_1 = p_1 \times q_1 \times r_1 \times s_1, \text{ \&c.}$$

$$P_2 = p_2 \times q_2 \times r_2 \times s_2, \text{ \&c.}$$

$$\dots\dots\dots$$

$$P_x = p_x \times q_x \times r_x \times s_x, \text{ \&c.}$$

and the value of an annuity of L.1 on the joint lives is $\Sigma v^x P_x$, from $x=1$ to $x=$ the number which renders any one of the probabilities p, q, r, s , &c. nothing.

58. The value of an annuity on the survivor of any num-

ber of given lives, that is, to continue so long as any one of them exists, is thus found. The probability that A will be alive at the end of the x th year being p_x , the probability that he will not be alive at the end of that time is $1-p_x$. The probability that *all* the lives will be extinct at the end of the x th year is therefore

$$(1-p_x)(1-q_x)(1-r_x)(1-s_x), \&c.$$

and the probability that they will not all be extinct, or that at least one of them will be in being, is

$$1-(1-p_x)(1-q_x)(1-r_x)(1-s_x), \&c.$$

which becomes by multiplication

$$p_x + q_x + r_x + s_x + \&c.$$

$$-p_x q_x - p_x r_x - \dots - q_x r_x - q_x s_x - \dots - r_x s_x - \&c.$$

$$+ p_x q_x r_x + p_x q_x s_x + \dots + q_x r_x s_x + \&c.$$

$$- p_x q_x r_x s_x - \&c.$$

$$+ \&c.$$

Multiplying each of the terms by v^x , and taking the sums of the respective products from $x=1$, and observing that $\Sigma v^x p_x q_x$ is the value of the annuity on the joint lives of A and B, $\Sigma v^x p_x q_x r_x$ that on the joint lives of A, B, and C, and so on, we have this rule:—

The value of an annuity on the survivor of any number of lives is equal to the sum of the annuities on each of the lives, minus the sum of the annuities on each pair of joint lives, plus the sum of the annuities on the joint lives taken by threes, and so on. When there are only two lives, the value of the annuity on the life of the survivor becomes

$$\Sigma v^x p_x + \Sigma v^x q_x - \Sigma v^x p_x q_x.$$

59. Let V denote the value of an assurance on the life of A, or the present worth of L.1 to be received at the end of the year in which A shall die. In respect of any year, the x th, after the present, the probability of A dying in the

course of that year is $p_{x-1} - p_x$. For let u be the probability that a life $x-1$ years older than A will live over one year, then $1-u$ is the probability of a life of that age not living over one year; therefore p_{x-1} being the probability of A living over $x-1$ years, $p_{x-1}(1-u)$ is the chance of his living over $x-1$ years, and dying in the following year (7). But $p_{x-1}(1-u) = p_{x-1} - p_x$; and by (53), $p_{x-1}u = p_x$; therefore $p_{x-1} - p_x$ is the chance that A will survive $x-1$ years and not survive x years. Now v^x is the value of L.1 to be received certainly at the end of the x th year; therefore in respect of the x th year the value of the expectation is $v_x(p_{x-1} - p_x)$; whence we have for the value of the assurance

$$V = \sum v^x (p_{x-1} - p_x),$$

from $x=1$ to $x=$ the number which makes $p=0$. Now, if we observe that $p_0=1$, and $\sum v^x p_{x-1} = v \sum v^{x-1} p_{x-1}$, it will be obvious that $\sum v^x p_{x-1} = v(1 + \sum v^x p_x)$; whence, denoting $\sum v^x p_x$ by A, (A being as in (54) the value of the annuity on the given life), we have

$$V = v(1 + A) - A; \text{ or } V = v - (1-v)A.$$

60. The values of assurances on joint lives, (that is, to be paid at the end of the year in which any one of the lives shall fail), or on the survivor of any number of joint lives, are calculated from the corresponding annuities by means of the same formula. Thus, let A' be the value of an annuity of L.1 on any number of joint lives, and V' the value of an assurance of L.1 on the same joint lives, then $V' = v - (1-v)A'$. If A'' be the annuity, and V'' the assurance on the life of the survivor of any number of given lives, we have still $V'' = v - (1-v)A''$.

61. Assurances on lives are usually paid not in single payments, but by equal yearly payments, the first being made

at the time the contract is entered into, and the succeeding ones at the end of each future year during the life of the assured. The present value of the sum which the assured contracts to pay is therefore equal to the first payment added to the value of an annuity of the same amount on his life; and if the assurance is made on terms of mathematical equality, this sum must be precisely equal to the value of the assurance in a single payment. Therefore, if y denote the amount of the yearly payment, we have the equation $y(1+A)=V$; whence $y=V\div(1+A)$.

62. The value of a temporary assurance for n years, that is, of an assurance to be paid only in the event of the individual dying before the end of n years is thus found. Let V be the present value of L.1, to be paid on the death of a person now aged y years, and V_n the present value of L.1, to be paid on the death of a person now aged $y+n$ years. At the end of n years from the present time, the value of L.1 assured on the life of a person now aged y years will be V_n , if he be then living. But the present value of L.1 to be received certainly at the end of n years is v^n ; and the probability that the life will continue n years is p_n ; therefore the present value of V_n , subject to the contingency of the life continuing n years, is $v^n p_n V_n$. If, therefore, we subtract this from V , we shall have the value of the temporary assurance in a single payment, namely $V-v^n p_n V_n$.

The equivalent annual premium is found by observing, that as the first payment is made immediately, and n payments are to be made in all, the value of all the premiums after the first is that of a temporary annuity of the same amount for $n-1$ years. Denoting therefore the annual premium by u , and the value of a temporary annuity for $n-1$ years by $A^{(tn')}$, the value of all the premiums is

$u + uA^{(tn')} = u(1 +)A^{(tn')}$; and we have consequently $u(1 + A^{(tn')}) = V - v^n p_n V_n$, whence

$$u = \frac{V - v^n p_n V_n}{1 + A^{(tn')}}.$$

63. The following question is of frequent occurrence. Required the present value of a sum of money to be received at the end of the year in which A dies, provided he die while B is living.

Let the sum be L.1, $W =$ its present value, $p_x =$ the probability of A living over x years, and $q_x =$ the probability of B living over x years. The chance of receiving the sum at the end of any given year, the x th, depends on two contingencies; 1. A may die in the course of that year, and B live over it; 2. A and B may both die in that year, A dying first. The probability of A dying in the x th year has been shewn (59) to be $p_{x-1} - p_x$; whence (7) the probability of the first contingency is $(p_{x-1} - p_x)q_x$. The probability that A and B will both die in the x th year is $(p_{x-1} - p_x)(q_{x-1} - q_x)$; and for so short a period as one year, it may be considered an even chance whether A or B will die first, whatever be the difference of their ages; therefore the probability in respect of the second contingency is $\frac{1}{2}(p_{x-1} - p_x)(q_{x-1} - q_x)$. Hence the whole probability of the sum being received at the end of the x th year, is $(p_{x-1} - p_x)q_x + \frac{1}{2}(p_{x-1} - p_x)(q_{x-1} - q_x) = \frac{1}{2}(p_{x-1} - p_x)(q_{x+1} + q_x)$, which being developed, and multiplied by v^x , becomes

$$\frac{1}{2}v^x(p_{x-1}q_{x-1} + p_{x-1}q_x - p_xq_{x-1} - p_xq_x),$$

and the sum of all the values of this expression from $x=1$, gives the value of W .

It has been already shewn (59) that $\Sigma v^x(p_{x-1} - p_x) = v - (1 - v)A$, where $A =$ the annuity on the life of A. In like

manner, if we denote by \overline{AB} the value of an annuity on the joint lives of A and B, we shall have $\Sigma v^x (p_{x-1} q_{x-1} - p^x q^x) = v - (1-v)\overline{AB}$, which is the value of an assurance to be paid on the death of the first dying. Assume p' such that $p'_x = p'_1 p_{x-1}$, then p'_x is evidently the probability that an individual A' one year *younger* than A, will live over x years (53), and $\Sigma v^x p_{x-1} q_x = \frac{1}{p'_1} \Sigma v^x p'_x q_x = \frac{1}{p'_1} \overline{A'B}$; denoting by $\overline{A'B}$ the value of an annuity on the joint lives of A' and B. Again, let $q'_x = q'_1 q_{x-1}$, then q'_x is the probability that B', who is one year *younger* than B, will live over x years, and $\Sigma v^x p^x q_{x-1} = \frac{1}{q'_1} \Sigma v^x p^x q'_x = \overline{AB'}$; denoting by $\overline{AB'}$ the value of an annuity on the joint lives of A and B'. Collecting the different terms, we have therefore

$$W = \frac{1}{2} \left\{ v - (1-v)\overline{AB} + \frac{1}{p'_1} \overline{A'B} - \frac{1}{q'_1} \overline{AB'} \right\}, \text{ whence } W$$

is easily computed from tables of annuities on joint lives.

If A and B are both of the same age, the two last terms destroy each other, and W is equal to $\frac{1}{2}$ the value of L.1, to be paid on the failure of the joint lives, as it evidently ought to be, since there is in this case the same chance of A dying before B as of B dying before A.

The formula gives the value of L.1 in a single payment; the equivalent yearly payment is W divided by $1 + \overline{AB}$, for the contract ceases on the failure of the joint lives by the death of either.

It would be easy to extend the formula to the case of an assurance to be paid on the contingency of the failure of any number of lives during the continuance of any number of other lives, or of an assurance to continue only during a stated time; but as it is not our purpose to give solutions

of the various problems of this kind which may occur in practice, but merely to shew the manner in which the general principles of the theory are applied to them, we shall not pursue the subject farther, but refer the reader to the article ANNUITIES, and to the standard works of Baily¹ and Milne,² in which it is treated in detail.

¹ *The Doctrine of Life Annuities and Assurances analytically investigated and practically explained, &c.* By Francis Baily. London, 1813. This work is now out of print, but a French translation of it has recently been published at Paris.

² *A Treatise on the Valuation of Annuities and Assurances on Lives and Survivorships, &c.* By Joshua Milne. London, 1815.

SECTION VII.

OF THE APPLICATION OF THE THEORY OF PROBABILITY
TO TESTIMONY, AND TO THE DECISIONS OF JURIES AND
TRIBUNALS.

64. The case of a witness making an assertion may be represented by an urn containing balls of two colours, the ratio of the number of one colour to that of the other being unknown, but presumed from the result of a number of experiments, which consist in drawing a ball at random, and replacing it in the urn after each trial. A true assertion being represented by a ball of one colour, and a false one by a ball of the other, it follows from the theorem in (51), that if a witness has made $m+n$ assertions, of which m are true and n false, the probability of a future assertion being true is $\frac{m+1}{m+n+2}$, and that of its being false $\frac{n+1}{m+n+2}$. Let the first of these fractions be represented by v , and the second by w , then v is the measure of the veracity of the individual, or the probability of his speaking the truth, and w the opposite probability, since $v+w=1$. In general, the existing data are insufficient to enable us to determine the numerical values of v and w in this manner; and therefore in applying the formulæ to particular cases, we must assign arbitrary values to these quantities, founded on previous knowledge of the moral character of the individual, or on

some notions, more or less sanctioned by experience, of the relative number of true and false statements made by men in general, placed in similar circumstances.

65. Having assumed v and w , let us suppose a witness to testify that an event has taken place, the *a priori* probability of which is p , and let it be proposed to determine the probability of the event after the testimony. In this case the event observed (E) is the assertion of the witness, and two hypotheses only can be made respecting its cause; 1st, that the event testified really took place; and 2d, that it did not. On the first hypothesis the witness has spoken the truth, the probability of which is v ; and an event has occurred of which the probability is p ; therefore (7) the probability (P_1) of the coincidence is vp . On the second hypothesis, the witness has testified falsely, the probability of which is w ; and the event attested did not happen, the probability of which is q ; therefore the probability (P_2) of the coincidence is wq . Hence, by the formula (47) $\pi_1 = P_1 \div \Sigma P_i$ the probability (π_1 , of the first hypothesis becomes

$$\frac{vp}{vp+wq}, \text{ and the probability } (\pi_2) \text{ of the second } \frac{wq}{vp+wq}$$

The sum of these two probabilities is unit, a condition which ought evidently to be fulfilled, since no other hypothesis can be made, and consequently one or other of the two must be true. It is to be observed, that these values of π_1 and π_2 are the respective probabilities, after the testimony has been given, that the event attested took place, and that it did not.

$$\begin{aligned} \text{Since } \pi_1 &= \frac{vp}{vp+wq}, \text{ we have } \pi_1 - p = \frac{p(v-vp-wq)}{vp+wq} \\ &= \frac{p\{v(1-p)-wq\}}{vp-wq} = \frac{p(vq-wq)}{vp-wq}; \text{ but } v-w=v-1+v= \end{aligned}$$

$2v-1$, therefore $\pi_1 - p = \frac{p(2v-1)q}{vp+wq}$. This fraction being positive or negative, according as $2v-1$ is greater or less than unity, or as v is greater or less than $\frac{1}{2}$, it follows that if $v > \frac{1}{2}$, then $\pi_1 > p$; that is to say, the probability of the event after the testimony is greater than its *a priori* probability when the veracity of the witness is greater than $\frac{1}{2}$. On the contrary, if the veracity of the witness is less than $\frac{1}{2}$, the effect of the testimony is to render the probability of the event less than its *a priori* probability.

66. If the event asserted by the witness be of such a nature that its occurrence is *a priori* extremely improbable, so that p is a very small fraction, and q consequently approaches nearly to unity, although at the same time the veracity of the witness be great, and measured by a fraction approaching to unity, the value of π_1 becomes nearly equal to $p \div w$, (for on this supposition $p + wq \div v$ is nearly equal to w). But it is obvious, that however great the improbability of a witness giving false testimony may be supposed, the improbability of a physical event may be any number of times greater; in other words, however small a value may be given to w , the value of p may still be any number of times smaller; so that notwithstanding the veracity of the witness, the probability of the event after the testimony, namely $\pi_1 = p \div w$ may be less than any assignable quantity. On this principle mankind do not easily give credence to a witness asserting a very extraordinary or improbable event. The odds against the occurrence of the event may be so great, that the testimony of no single witness, however respectable his character, would suffice to induce belief.

67. In the case of the character of a witness being alto-

gether unknown, we may suppose v to have all possible values within certain limits, and to find the value of π_1 by integrating the fraction $\int \pi_1 dx$ between those limits. Since $\pi_1 = \frac{vp}{vp+wq}$, we have $\int \pi_1 dx = \int \frac{pvdv}{vp+wq}$, which on substituting $1-v$ and $1-p$ for w and q respectively, becomes

$\int \frac{pvdv}{1-p+(2p-1)v}$, the integral of which is

$$p \left\{ \frac{v}{2p-1} - \frac{1-p}{(2p-1)^2} \log. (1-p+(2p-1)v) \right\} + C,$$

C being a constant, the value of which will be determined from the assumed limits. If v be supposed to vary between the limits $v=0$ and $v=1$, then

$$\int \pi_1 dx = \frac{p}{2p-1} \left\{ 1 - \frac{1-p}{2p-1} \log. \frac{p}{1-p} \right\};$$

and if we assume $p=\frac{3}{4}$, we have $\int \pi_1 dx = \frac{1}{2}(1-\frac{1}{2} \log. 3)$, which, since the logarithm is the Napierian logarithm, and Nap. log. 3=1.0986, becomes $\frac{1}{2} \times .4507 = .676$, or nearly $\frac{2}{3}$. Whence we see, that on this hypothesis the probability of the event is diminished in consequence of the testimony.

68. The credit due to the testimony of a witness depends not merely on his good faith, but also on the probability that he is not himself deceived with respect to the event he asserts. The chances of a witness being deceived through credulity or ignorance are much more numerous in general than the chances of intentional fraud; and this must be the case more particularly when the event is of such a nature that it may happen in various ways which may be mistaken one for another: as for instance, in the case of a lottery ticket being drawn, and the witness asserting that it bears a particular number, which might with equal probability be any other number on the wheel. The following

question will illustrate the method of applying the calculus when a distinction is made between these sources of error.

An urn contains s balls, of which a_1 are marked A_1 , a_2 marked A_2 , a_n marked A_n . A ball having been drawn at random, a witness of the drawing affirms that the ball drawn is marked A_m ; required the probability of the testimony being true.

Here we have $s = a_1 + a_2 + a_3 + \dots + a_n$ (n being the number of the different indices or sorts of balls); so that if we make $p_1 = a_1 \div s$, $p_2 = a_2 \div s$, $p_m = a_m \div s$, then p_1 is the *a priori* probability that the ball drawn is of the class marked A_1 , p_2 the probability that it belongs to the class whose index is A_2 , and so on. It is evident that n different hypotheses may be made respecting the index of the ball which has been drawn, for it may belong to any one of the different classes $A_1, A_2 \dots A_n$. Let the probabilities of these hypotheses be respectively $\varpi_1, \varpi_2 \dots \varpi_n$, (that is, in respect of any particular index i , ϖ_i is the probability after the assertion that the ball drawn is marked A_i); and let the probabilities of the assertion on each of these hypotheses be respectively $P_1, P_2 \dots P_n$, (that is, if the ball drawn be marked A_i , then P_i is the probability the witness will assert it to be marked A_m). Lastly, let v be the veracity of the witness, and u the probability that he has not been deceived.

(1.) Let us first consider the hypothesis that the ball drawn is marked A_m , and consequently that the assertion is true. In order to find P_m , the probability of the assertion being made, there are four cases to be considered. 1st, we may suppose the witness is not deceived himself (u), and that he speaks the truth (v). The probability of the assertion in this case is uv . 2d, The witness knows the truth, but intends to deceive, or testifies falsely. In this case the proba-

bility of the assertion being made, on the hypothesis under consideration, is 0. 3d, The witness has been deceived himself, but intends to speak the truth. In this case also the probability of the assertion being made is 0. 4th, The witness has been deceived himself, and intends to deceive. In this case the assertion might be made; and to find the probability of its being made we have to consider, that since the witness has been deceived, he must have supposed some other index than A_m to have been drawn; and since he intends to deceive, he must assert some other index to be drawn than that which he supposes to be drawn. Setting aside, therefore, the index which he supposes to have been drawn, there remain $n-1$ others, any one of which he is as likely to name as any other. The probability, therefore, of his naming A_m when he intends to deceive is $1 \div (n-1)$. Hence the probability of the assertion in this case is compounded of the probabilities of three simple events, as follows: 1. Probability the witness is deceived $= (1-u)$; 2. Probability he intends to deceive $= (1-v)$; 3. Probability he names $A_m = 1 : (n-1)$. The probability of the assertion is therefore in this case $= (1-u)(1-v) \div (n-1)$. Adding this to the probability found in the first case, we have P_m , the whole probability of the assertion being made on the hypothesis that the index of the ball drawn was A_m , namely

$$P_m = uv + \frac{(1-u)(1-v)}{n-1}.$$

(2.) Let us now consider one of the remaining hypotheses, and suppose that the ball actually drawn was marked A_b and not A_m , as attested by the witness. As before, there are four possible cases for consideration. 1st, The witness knows the fact and speaks the truth. In this case the assertion could not be made, or its probability is 0. 2d, The

witness knows the fact, and intends to deceive. In this case the probability of his asserting A_m to be drawn is compounded of the probability that he is not deceived (u), the probability that he testifies falsely ($1-v$), and the probability that, knowing the index A_i to be drawn, he selects A_m from among the $n-1$ which remain after rejecting A_n ($1 \div (n-1)$). The probability of the assertion being made in this case is therefore $u(1-v) \div (n-1)$. 3d, The witness is deceived, and intends to speak the truth. By reasoning as in the last case, it is easy to see that the probability of the assertion being made in this case is $v(1-u) \div (n-1)$. 4th, The witness is deceived, and intends to deceive. The probability of the assertion being made in this case will be found by considering, that as the witness is himself deceived, he must suppose some particular index to be drawn different from A_n (which is drawn by hypothesis), for instance A_e , the probability of which is $1 \div (n-1)$; and intending to deceive, he must fix on some index different from A_e , which he supposes to be drawn; and he announces A_m , the probability of which selection is also $1 \div (n-1)$. The probability, therefore, that the witness supposes A_e to be drawn, and announces A_m , is $1 \div (n-1)^2$. But it is evident, that whatever can be affirmed with respect to the particular index A_e , may be affirmed with equal truth of every one of the other n indexes, excepting A_n which is actually drawn, (since by hypothesis the witness is deceived), and A_m , which he announces, (since by hypothesis he lies). There are therefore $n-2$ different ways in which he may at the same time be deceived, and intend to deceive, and announce A_m ; consequently the probability of this announcement in any of these ways is $(n-2) \div (n-1)^2$. Multiplying this into the probability of his being deceived ($1-u$), and the probability

of his giving false testimony ($1-v$), the probability of the assertion in this case becomes $\frac{(1-u)(1-v)(n-2)}{(n-1)^2}$. Hence the whole probability of the assertion, in all the cases included in the hypothesis that the ball actually drawn was marked A_v is

$$P_i = \frac{u(1-v)}{n-1} + \frac{(1-u)v}{n-1} + \frac{(1-u)(1-v)(n-2)}{(n-1)^2}.$$

As this expression will evidently be the probability of the assertion on any other of the $n-1$ hypotheses that the ball actually drawn was marked with an index different from A_m , the sum of the probabilities of the assertion on all these hypotheses is ΣP_i where i is successively each of the numbers 1, 2, 3, ..., n , excepting m .

We have now to find ϖ_m , the probability of the first hypothesis. Since the hypotheses, in the present question, are not all equally probable *a priori*, we must have recourse to the formula (49) $\varpi_i = \lambda_i P_i \div \Sigma \lambda_i P_i$, and consequently in the present case we have

$$\varpi_m = \frac{\lambda_m P_m}{\lambda_m P_m + \Sigma \lambda_i P_i},$$

the sign of summation Σ including every value of i from 0 to n , excepting $i=m$. Now the value of P_i being the same in respect of each of the hypotheses which suppose the assertion untrue, $\Sigma \lambda_i P_i = P_i \Sigma \lambda_i$; and the sum of all the values of λ_i from $i=0$ to $i=n$ being 1, on excluding $\lambda_m = a_m \div s$, we have $\Sigma \lambda_i = (s - a_m) \div s$. Substituting this, together with the values of P_m and P_i as above found, and making $u' = 1-u$, $v' = 1-v$, the formula becomes, after the proper reduction,

$$\varpi_m = \frac{a_m \{ (n-1)uv + u'v' \}}{a_m \{ (n-1)uv + u'v' \} + (s-a_m) \{ u'v + uv' + \frac{n-2}{n-1} u'v' \}};$$

which is the probability of the hypothesis that a ball marked A_m was drawn, or that the testimony is true.

When there are no two balls in the urn having the same index, the numbers $a_1, a_2, a_3, \&c.$ become each $=1$, and $s=n$. In this case the formula gives

$$\pi_m = \frac{(n-1)uv + u'v'}{(n-1)uv + u'v' + (n-1)(u'v + uv') + (n-2)u'v'},$$

which, on observing that $uv + u'v + uv' + u'v' = 1$, becomes by reduction

$$\pi_m = uv + \frac{(1-u)(1-v)}{n-1}.$$

This is the probability of the truth of the testimony of a witness, who affirms that the number m is drawn from an urn which contains n balls, numbered $1, 2, 3 \dots n$. It is obvious, that when u and v are fractions approaching to unity, and n is a considerable number, the second term becomes very small, and may be neglected. The probability then becomes simply $\pi_m = uv$.

69. We now proceed to consider the probability of an event attested by several witnesses; and first let us suppose the witnesses to agree in their testimony. The measures of the veracity of the several witnesses being respectively $v_1, v_2, v_3, \&c.$, and the *a priori* probability of the event being p , we have by (58) for its probability after the testimony of the first witness,

$$\pi_1 = \frac{v_1 p}{v_1 p + (1-v_1)(1-p)}.$$

In order to find the probability of the event after the second witness gives his testimony, we may suppose the *a priori* probability to be changed from p to π_1 by the testimony of the first witness, and the same formula gives

$$\pi_2 = \frac{v_2 \pi_1}{v_2 \pi_1 + (1-v_2)(1-\pi_1)} = \frac{v_1 v_2 p}{v_1 v_2 p + (1-v_1)(1-v_2)(1-p)}.$$

Let a third witness now come forward, and give testimony in favour of the same event. Its probability after his testimony will become in like manner

$$\pi_3 = \frac{v_3 \pi_2}{v_3 \pi_2 + (1-v_3)(1-\pi_2)} = \frac{v_1 v_2 v_3 p}{v_1 v_2 v_3 p + (1-v_1)(1-v_2)(1-v_3)(1-p)}.$$

In general, let π_x be the probability of an event after it has been attested by x witnesses, and let v_x be the veracity of the last witness, then π_{x-1} being the probability of the event after $x-1$ eyewitnesses have each testified in its favour, we have

$$\pi_x = \frac{v_1 v_2 \dots v_x p}{v_1 v_2 \dots v_x p + (1-v_1)(1-v_2) \dots (1-v_x)(1-p)}.$$

If we suppose the witnesses all equally credible, or that $v_1 = v_2 = v_3 \dots = v_x$, this becomes

$$\pi_x = \frac{v^x p}{v^x p + (1-v)^x (1-p)} = 1 + \left(\frac{1-v}{v} \right)^x \frac{1-p}{p}.$$

Now, if $v = \frac{1}{2}$, then $(1-v) \div v = 1$, and $\pi_x = p$; whence it appears that the probability of an event is not increased by the testimony of any number of witnesses, when the veracity of each is only $\frac{1}{2}$; but when v is greater than $\frac{1}{2}$, the event becomes more probable as the number of witnesses is greater, and when v is a considerable fraction, its probability increases very rapidly with the number of witnesses.

70. When the values of v and p are given, that of x in the last formula may be found so as to render π_x of any given value. Hence we may find the number of witnesses required to make it an even wager, whether an event exceedingly improbable, and in favour of which they give una-

nimous testimony, has happened or not. For example, let the odds against the event be a million million to one,

that is, let $p = \frac{1}{1,000,000,000,001} = \frac{1}{10^{12}+1}$, and let v the

veracity of each witness be $\frac{9}{10}$. In order that πx may equal $\frac{1}{2}$,

we must have $\left(\frac{1-v}{v}\right)^x \frac{1-p}{p} = 1$. Now $\frac{1-v}{v} = \frac{1}{9}$, and

$\frac{1-p}{p} = 10^{12}$, therefore $\left(\frac{1}{9}\right)^x \times 10^{12} = 1$; whence $x \log \frac{1}{9} =$

$\log \frac{1}{10^{12}}$ or $x \log 9 = 12$ and therefore $x = \frac{12}{.95424} = 12.6$

nearly, so that 13 independent witnesses would suffice to render it more probable that the event really took place than that it did not.

This example is given by Mr. Babbage, (Ninth Bridgewater Treatise, Note E), with a view to shew the fallacy of Hume's celebrated argument respecting miracles. What the example proves is simply this, that if we suppose an urn to contain a million million of white balls, and only one black ball, and that on a ball being drawn at random from the urn, thirteen eyewitnesses of the drawing, each of whom makes only one false statement in ten, without collusion, and independently of each other, affirm to A, who was not present at the drawing, that the ball drawn was black, then A would have rather a stronger reason for believing than for disbelieving the testimony. But it is sufficiently obvious, that the event attested in this case, though exceedingly improbable *a priori*, cannot be regarded as in any way miraculous. On the contrary, the black ball might be drawn with the same facility, and was *a priori* as likely to be drawn,

as any other *specified* ball in the urn. Let it be granted that an event is within the range of fortuitous occurrence, and that there exists a single chance in its favour out of any number of millions of chances, it may then happen in any one trial; nay, a number of trials may be assigned, such that its non-occurrence would be many times more improbable than the contrary.

71. Let us next consider the case of a number of witnesses contradicting each other. If the first witness announces an event of which the probability is p , then the probability, after the testimony, of its having happened is π_1 , and the probability that it has not happened $1-\pi_1$. Suppose a second witness now to appear, and testify that the event has not happened, and let the probability of the truth of his testimony be denoted by π'_2 ; then $1-\pi_1$ being the probability before his testimony was given that what he asserts is true, and v_2 being the measure of his veracity, we

have, as in (69), $\pi'_2 = \frac{v_2(1-\pi_1)}{v_2(1-\pi_1) + (1-v_2)\pi_1}$ whence, since

$$\pi_1 = \frac{v_1 p}{v_1 p + (1-v_1)(1-p)}, \text{ there results}$$

$$\pi'_2 = \frac{(1-v_1)v_2(1-p)}{(1-v_1)v_2(1-p) + v_1(1-v_2)p}$$

for the measure of the probability that the event has not happened. The probability that it has happened is therefore $1-\pi'_2$, and accordingly if π'_2 be less than $\frac{1}{2}$, there is a stronger reason for believing that the event happened than that it did not. The method of forming the expression for the probability of the event, after it has been attested or denied by a third witness, or any number of successive witnesses, is obvious.

If we suppose the values of v_1 and v_2 to be equal, the expression becomes $\pi'_2 = 1 - p$, which is the *a priori* probability that the event did not happen. It is obvious that this must be the case, inasmuch as two contradictory testimonies of equal weight neutralize each other. In general, the probability of an event which is affirmed by m witnesses, and denied by n witnesses, all equally credible, is the same as that of an event which is affirmed by $m - n$ witnesses who agree in their testimony.

72. When a relation has been transmitted through a series of narrators, of whom the first only has a direct knowledge of the event, and each of the others derives his knowledge from the relation of the preceding, the probability of the event is diminished by every succeeding relation. In order to obtain a general expression for the probability of traditionary testimony, we may take the event considered in (68), namely the extraction of a ball marked A_m from an urn containing s balls, of which a_1 are marked A_1 , a_2 marked A_2 , ..., a_n marked A_n , there being in all n different indices. Now suppose the relation to have passed through a chain of narrators, T, T_1, T_2, \dots, T_x , in number $x + 1$, of whom the first only was an eyewitness of the event, each of the others receiving his knowledge of it from the one preceding him, and communicating it in his turn to the succeeding, the question is to determine the probability that a ball marked A_m was drawn, after this event has been narrated by T_x , the last witness of the series.

73. In order to apply the general formula of (68) to this case, it is necessary to remark that the event observed is the attestation of T_x of his having been informed by T_{x-1} that the ball drawn from the urn, the drawing of which was seen by T , was marked A_m . There are n different hypotheses

respecting the index of the ball actually drawn, but it is only necessary to consider two of them, namely, the hypothesis that the ball actually drawn was marked A_m , and any one of the other hypotheses which consist in supposing that a ball with a different index from A_m was drawn, for example A_i . Let the probability of the attestation, on the hypothesis that the index of the ball drawn was A_m , be denoted by y_x , and its probability on the hypothesis that the index was A_i by y'_x , (y_x and y'_x corresponding to P_m and P_i in (68), which express the same probabilities in respect of the eyewitness T), then by (68), the probability of the hypothesis that A_m was drawn is

$$w_m = \frac{\lambda_m y_x}{\lambda_m y_x + \sum \lambda_i y'_x}.$$

But since y'_x is the same for all the hypotheses that the index drawn was different from A_m , $\sum \lambda_i y'_x = y'_x \sum \lambda_i$; and by (68) $\sum \lambda_i = (s - a_m) \div s$, and $\lambda_m = a_m \div s$, therefore

$$w_m = \frac{a_m y_x}{a_m y_x + (s - a_m) y'_x}.$$

We have now to find y_x and y'_x in terms of x . Let v , v_1 , v_2 , &c. be the respective probabilities of T , T_1 , T_2 , &c. speaking the truth, then the probability of T_x speaking the truth is v_x , and the probability that he does not $1 - v_x$, whether because he is dishonest, and intends to deceive, or because he has mistaken the statement of the preceding witness. Now there are two ways in which it may happen that A_m is announced by T_x . First, if he speaks the truth, and has been informed by the preceding narrator T_{x-1} that A_m was the index drawn; secondly, if he lies, and has been informed by T_{x-1} that a different index from A_m was drawn. Assuming y_{x-1} to have the same signification with respect to T_{x-1} that y_x has been assumed to have with respect to T_x , (that is to say,

the probability of the assertion being made by T_{x-1} on the hypothesis that the ball actually drawn was A_m), the probability of the first of these combinations is $v_x y_{x-1}$. With respect to the second case, it is to be observed, that if T_x announces a different index from that which has been announced to him by T_{x-1} , the chance of his announcing A_m out of $n-1$ indexes different from that announced by T_{x-1} is $1 \div (n-1)$; and on multiplying this by the probability $1-v_x$ that the testimony of T_x is false, and by the probability $1-y_{x-1}$ that T_{x-1} has announced a different index from A_m , we have, for the probability of the second combination $(1-v_x)(1-y_{x-1}) \div (n-1)$. The whole probability of T_x testifying that A_m was drawn, is therefore, on the first hypothesis, given by the equation,

$$y_x = v_x y_{x-1} + (1-v_x)(1-y_{x-1}) \div (n-1).$$

This is an equation of finite differences of the first order, the complete integral of which is

$$y_x = \frac{1}{n} + \frac{C(nr_1-1)(nr_2-1)\dots(nr_{x-1}-1)(nr_x-1)^1}{(n-1)^x}.$$

In order to determine the arbitrary constant C , it is to be observed, that, since $y_1, y_2 \dots y_x$, as well as $v_1, v_2 \dots v_x$ apply to the narrators $T_1, T_2 \dots T_x$ respectively, if we suppose $x=0$ the resulting value of the integral will be the probability that A_m was announced by the eyewitness T , on the hypothesis that A_m was actually drawn. Let this probability be P_m ; then the equation becomes $P_m = C + 1 \div n$ whence $C = (nP_m - 1) \div n$. If, therefore, we make.

$$X = \frac{(nr_1-1)(nr_2-1)\dots(nr_{x-1}-1)(nr_x-1)}{(n-1)^x}$$

¹ This is easily verified; for on changing x into $x-1$ in the integral, and forming the expression $v_x y_{x-1} + (1-v_x)(1-y_{x-1}) \div (n-1)$, there results an identical equation.

we obtain, on the first hypothesis, for any value of x ,

$$y_x = \{1 + (nP_m - 1)X\} \div n$$

In the same manner we find the probability y'_x of the testimony given by T_x , on the hypothesis that the ball 'actually drawn was A_i . The probability of the event, after being testified by the eyewitness, being on this hypothesis P_o we have,

$$y'_x = \{1 + (nP_i - 1)X\} \div n.$$

Substituting these values of y_x and y'_x in the expression above found for ϖ_m , we obtain for the probability of the event observed by T , and narrated by T_x , the narration having passed from one to another in the manner supposed,

$$\varpi_m = \frac{a_m \{1 + n(P_m - 1)X\}}{a_m \{1 + (nP_m - 1)X\} + (s - a_m) \{1 + (nP_i - 1)X\}}.$$

74. Since $\frac{nv_x - 1}{n - 1} = v_x - \frac{1 - v_x}{n - 1}$, and since v_x is always less

than unity, and n always greater than unity, each of the terms of the series represented by X , whether positive or negative, is a proper fraction, whence the value of X becomes smaller and smaller as x increases. Suppose x infinite, then $X = 0$, and $\varpi_m = a_m \div s$, which is the *a priori* probability of the event. Hence we see that the probability of an event transmitted through a series of traditionary evidence becomes weaker at every step, and ultimately equal to the simple probability of the event, independent of any testimony.

75. When the urn is supposed to contain only n balls, each having a different index, the expression for ϖ_m is greatly simplified; for, in this case, $a_m = 1$, $s = n$; therefore, (since $P_m + (n - 1)P_i = 1$) the denominator becomes n , and we have consequently $\varpi_m = \{1 + (nP_m - 1)X\} \div n$, which coincides

with the value of y_x found above, that is to say, with the probability of the event being testified by T_x on the hypothesis that it actually happened. Laplace, in solving this particular case of the problem, (p. 456) assumes that the probabilities here denoted by y_x and π_m are identical. They are, however, as is evident from the above analysis, quite distinct in their nature, and their values are only equal in the particular case in which $s—a_m$ is to a_m in the ratio of $m—1$ to 1. (Poisson, p. 112.)

76. The question of determining the probability that the verdict of a jury is correct, is precisely analogous to that of finding the probability of an event attested by one or more witnesses. Let us first take the case of a single juror, and assume $u =$ the probability that the juror gives a correct verdict, (that is, correct in respect of the facts), and $p =$ the probability that the accused is guilty before being put on his trial. Suppose the verdict guilty to be returned; two hypotheses may be made respecting the cause of the verdict, first, that the accused is guilty; secondly, that he is innocent. On the first hypothesis, the accused will be condemned if the juror gives a right verdict, the probability of which is u . On the second hypothesis, the accused will be condemned if the juror gives a wrong verdict, the probability of which is $1—u$. But the *a priori* probabilities of these causes (the guilt or innocence of the accused) being respectively p and $1—p$, we have by (49)

$$\pi_1 = \frac{up}{up + (1-u)(1-p)}, \quad \pi_2 = \frac{(1-u)(1-p)}{up + (1-u)(1-p)},$$

π_1 being the probability of the first hypothesis, or the probability that the accused is guilty after the verdict has been given, and π_2 the probability resulting from the verdict that the accused is innocent.

77. Suppose the verdict not guilty to be given, and let π'_1 and π'_2 be the probabilities after the verdict of the two hypotheses. On the first hypothesis, namely, that the accused is guilty, this verdict will be given if the juror gives a wrong verdict, of which the probability is $1-u$; and on the second hypothesis, the verdict will be given if the juror gives a right verdict, of which the probability is u ; and the probabilities of these hypotheses before the verdict being respectively p and $1-p$ as before, we have

$$\pi'_1 = \frac{(1-u)p}{(1-u)p + u(1-p)}, \quad \pi'_2 = \frac{u(1-p)}{(1-u)p + u(1-p)}.$$

From the above value of π_1 , we obtain $\pi_1 - p = \frac{p(1-p)(2u-1)}{up + (1-u)(1-p)}$; a fraction which is positive or negative according as u is greater or less than $\frac{1}{2}$. Hence it appears that the guilt of the accused is only rendered more probable by the verdict guilty being pronounced, when the probability that the juror gives a correct verdict is greater than $\frac{1}{2}$. In like manner it is shewn that π'_1 (the presumption of the guilt of the accused after a verdict of acquittal), is greater than p when u is less than $\frac{1}{2}$.

78. The *a priori* probability of the condemnation of the accused before he is put on his trial is $up + (1-u)(1-p)$; for there are two ways in which this condemnation may take place; first, if the accused be guilty, and the juror give a correct verdict, the probability of which concurrence is up ; and, secondly, if the accused be innocent, and the juror give a wrong verdict, the probability of which is $(1-u)(1-p)$. Therefore, making c = the probability of a verdict of condemnation, we have $c = up + (1-u)(1-p)$; and for a verdict of acquittal, $1-c = (1-u)p + u(1-p)$.

79. Let us next suppose that after the verdict of the first

juror has been pronounced, the accused is put on his trial before a second juror, and let u_2 be the probability the second juror gives a correct verdict, and c_2 be the probability the accused will be pronounced guilty by him. After the verdict guilty has been pronounced by the first juror, the probability of the guilt of the accused is π_1 , and it is evident that c_2 will be found by substituting u_2 for u , and π_1 for p in the above value of c , whence $c_2 = u_2 \pi_1 + (1 - u_2)(1 - \pi_1)$. The probability of a verdict of condemnation by both jurors is cc_2 ; therefore, (observing that $\pi_1 = up \div c$), we have for this probability

$$cc_2 = uu_2p + (1 - u)(1 - u_2)(1 - p).$$

The probability of the guilt of the accused after a verdict of acquittal has been pronounced by the first juror being π'_1 , the probability of a verdict of acquittal being given by the second juror is $1 - c_2 = (1 - u_2)\pi'_1 + u_2(1 - \pi'_1)$; therefore, (observing that $\pi'_1 = (1 - u)p \div (1 - c)$), we have for the probability of a verdict of acquittal by both jurors

$$(1 - c)(1 - c_2) = (1 - u)(1 - u_2)p + uu_2(1 - p).$$

Adding the probability of a verdict of condemnation by both jurors, to that of acquittal by both, we have $uu_2 + (1 - u)(1 - u_2)$ for the probability of both giving the same verdict. This result is independent of p , and is evidently true *a priori*, inasmuch as there are two ways in which the same verdict may be given, namely, when both jurors are right, and when both are wrong.

The probability of acquittal by the second juror, after a verdict of guilty by the first, is $1 - c_2 = (1 - u_2)\pi_1 + u_2(1 - \pi_1)$; multiplying by c , and substituting for c and π_1 their values, we have for the probability of a verdict of guilty by the first, and not guilty by the second,

$$c(1 - c_2) = u(1 - u_2)p + (1 - u)u_2(1 - p).$$

In like manner, if the accused has been acquitted by the first juror, the presumption of his guilt becomes π'_1 , and the probability of a verdict guilty by the second is $c_2 = u_2 \pi'_1 + (1 - u_2)(1 - \pi'_1)$; therefore the probability of a verdict of not guilty by the first, and of guilty by the second is

$$(1 - c)c_2 = (1 - u)u_2p + u(1 - u_2)(1 - p).$$

The sum of these two expressions gives for the probability of a discordant verdict, $u(1 - u_2) + (1 - u)u_2$.

80. If we now suppose $u = u_2$, and make $1 - u = w$, the probability that the two jurors will agree in their verdict, whether they are both right or both wrong, is $u^2 + w^2$; and the probability of a discordant verdict $uw + uw = 2uw$. The sum of the two expressions is $u^2 + 2uw + w^2 = (u + w)^2$; and therefore the probabilities of the different cases are respectively given by the developement of the binomial $(u + w)^2$.

By pursuing this reasoning, it is easy to see that if there be any number h whatever of jurors, or voters on any question which admits only of simple affirmation or negation, all being supposed to possess the same integrity and knowledge, so that there is the same probability u of a correct decision in respect of each, the probabilities of the different cases are found by the developement of the binomial $(u + w)^h$. The probability of a correct verdict being pronounced unanimously is u^h ; of an erroneous one being pronounced unanimously is w^h ; and the probability that a correct verdict will be given by m of the jurors, and an erroneous one by n , is $Uu^m w^n$, where $U = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n}$.

81. The probability that the accused will be pronounced guilty by m jurors, and acquitted by n , on the supposition that the value of u is the same for each juror, is thus found. There are two ways in which this event may take place;

1st, if the accused be guilty (the probability of which is p), and m jurors decide correctly, and n wrongly (the probability of which is Uu^mw^n); the probability of the condemnation taking place in this way is therefore Uu^mw^np . 2d, If the accused be innocent (the probability of which is q) and n jurors decide rightly, and m wrongly (the probability of which is Uu^nw^mq); the probability of the event taking place in this way is therefore Uu^nw^mq . Let G therefore denote the whole probability of the verdict, and we have

$$G = U(u^mw^np + u^nw^mq).$$

Hence the probability that the accused will be condemned unanimously by a jury consisting of h jurors is $u^hp + w^hq$; and the probability that he will be unanimously acquitted $u^hq + w^hp$.

82. Suppose the accused to have been pronounced guilty by m jurors, and not guilty by n jurors, the probability of the verdict of the majority being correct is found from the formula in (49). Two hypotheses may be made: 1st, the accused is guilty; 2d, he is innocent. The probability P_1 of the observed event (the condemnation by m , and acquittal by n jurors) on the first hypothesis is Uu^mw^n ; and the *a priori* probabilities of the two hypotheses (or the probabilities denoted by λ_1 and λ_2 in (49), being p and q ; therefore if π_1 denote the probability of the verdict being correct, that is, the probability of the first hypothesis after the verdict has been pronounced, and π_2 the probability of its being wrong, we shall have (49)

$$\pi_1 = \frac{u^mw^np}{u^mw^np + u^nw^mq}, \quad \pi_2 = \frac{u^nw^mq}{u^mw^np + u^nw^mq}.$$

If the verdict has been pronounced unanimously, then $m=h$ and $n=0$, and the formulæ become

$$\pi_1 = \frac{u^h p}{u^h p + w^h q}, \quad \pi_2 = \frac{w^h q}{u^h p + w^h q}.$$

If $p=q=\frac{1}{2}$, and $m-n=i$, we have then

$$\pi_1 = \frac{u^{n+i} w^n}{u^{n+i} w^n + u^n w^{n+i}} = \frac{u^i}{u^i + w^i}.$$

But this is the probability of a verdict being correct which has been pronounced unanimously by i jurors; whence it follows that the probability of a decision rendered by a given majority being correct, is the same as that of a decision rendered unanimously by a jury equal in number to the difference between the majority and minority, and is therefore independent of the total number of jurors. This, however, is only true on the supposition that the value of u is known *a priori*; for if u be not absolutely known, the weight of the verdict depends on the ratio of the majority to the whole number of jurors. This is in accordance with common notions, for it will readily be admitted that a verdict given unanimously by a jury of 10 will be entitled to much more weight than one pronounced by a jury consisting of a large number, as 100, in which 55 are of one opinion, and 45 of the opposite. In this case, the opinion of the minority throws great doubt on the correctness of the verdict. It is to be observed, however, that the probability of a verdict being given by a small majority becomes less and less as the number of jurors is increased.

83. When the number who dissent from the opinion of the majority is unknown, and we merely know that the majority exceeds the minority by *at least* i jurors, the probability of the verdict being correct is found as follows. Suppose the verdict to be guilty. On the hypothesis that it is correct, the probability of the accused being found guilty by $h-x$, and not guilty by x jurors, is by the formula in (80),

$Uu^{h-x}w^x$. Now, if we give x successively all the values 0, 1, 2, ... n , where $n = \frac{1}{2}(h-i)$, and assume U_0 to denote the value of U when $x=0$, U_1 its value when $x=1$, and so on; and also make W = the probability of the accused being pronounced guilty by $h-n$ at least, we shall have

$$W = U_0 u^h + U_1 u^{h-1} w + U_2 u^{h-2} w^2 + \dots + U_n u^{h-n} w^n.$$

In like manner, if W' denote the probability of a verdict guilty by $h-n$ jurors at least, on the hypothesis that the accused is not guilty, we shall have,

$$W' = U_0 w^h + U_1 u w^{h-1} + U_2 u^2 w^{h-2} + \dots + U_n u^n w^{h-n};$$

whence, p and q being as above the *a priori* probabilities of the two hypotheses, the probability that the verdict guilty is correct, when pronounced by $h-n$ jurors at least, becomes $Wp \div (Wp + W'q)$.

84. It is evident that no application can be made of these formulæ without assigning arbitrary values to u and p , unless, indeed, we have data for determining their mean values from experience. With respect to p , we may assume, for the sake of shewing the general consequences of the formulæ, its value to be $\frac{1}{2}$; for it cannot well be supposed less than $\frac{1}{2}$, or that a person brought before a jury is more likely to be innocent than guilty; and if it much exceeds $\frac{1}{2}$ and approaches to unity, a verdict of guilty may be expected from any jury, however constituted. When a mean value of u cannot be determined from experience, the only way of obtaining numerical results, is to suppose u to have all possible values within given limits, and to integrate the equations between those limits. As it seems unreasonable to suppose that a juror is more likely to give a wrong verdict than a right one, we may assume that u cannot be less than $\frac{1}{2}$. Suppose, then, that u increases by infinitely small increments

from $u=\frac{1}{2}$ to $u=1$, and let it be proposed to determine the probability that a decision is correct when the accused has been pronounced guilty by m jurors, and not guilty by n . Here an infinite number of hypotheses may be made respecting the value of u , and we must therefore have recourse to the formulæ in (51.) Let $u=x$ be one of those hypotheses, $P_x=$ the probability on that hypothesis of the event observed (that is, of the accused being pronounced guilty by m , and not guilty by n jurors,) $\varpi_x=$ the probability of the assumed hypothesis, and $\Pi=$ the mean probability of the correctness of the verdict from all the hypotheses. By the formulæ in (81) we have

$$P_x = U \{x^m(1-x)^n p + x^n(1-x)^m (1-p)\},$$

and as all the hypotheses are supposed equally probable, we have (45) $\varpi_x = P_x \div \Sigma P_x$. But between the proposed limits $\Sigma P_x = p \int_{\frac{1}{2}}^1 x^m (1-x)^n dx + (1-p) \int_{\frac{1}{2}}^1 x^n (1-x)^m dx$; if, therefore, we make $p=\frac{1}{2}$, we shall have by reason of $\int_{\frac{1}{2}}^1 x^n (1-x)^m dx = \int_{\frac{1}{2}}^1 x^m (1-x)^n dx$, $\Sigma P_x = \frac{1}{2} \int_0^1 x^m (1-x)^n dx$, and therefore

$$\varpi_x = \frac{x^n (1-x)^n + x^n (1-x)^m}{\int_0^1 x^m (1-x)^n dx},$$

for the probability of the hypothesis. But (82) the probability on this hypothesis of the accused being guilty, is

$\frac{x^m (1-x)^n}{x^m (1-x)^n + x^n (1-x)^m}$; multiplying this by the probability of the hypothesis, ϖ_x , we obtain for the probability of the verdict being correct $x^m (1-x)^n \div \int_0^1 x^m (1-x)^n dx$; and, therefore, for the probability of the verdict being correct on all the hypotheses from $x=\frac{1}{2}$ to $x=1$,

$$\Pi = \frac{\int_{\frac{1}{2}}^1 x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx}.$$

Hence the probability that a verdict given by a majority m out of $m+n=h$ jurors is wrong, is

$$1-\Pi = \frac{\int_0^{\frac{1}{2}} x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx},$$

which, on effecting the integrations by the formula in (51) becomes after reduction

$$1-\Pi = \frac{1}{2^{h+1}} \left\{ 1 + \frac{h+1}{1} + \frac{(h+1)h}{1 \cdot 2} + \frac{(h+1)h(h-1)}{1 \cdot 2 \cdot 3} \dots \dots + \frac{(h+1)h(h-1) \dots \dots (h-n+2)}{1 \cdot 2 \cdot 3 \dots \dots n} \right\}.$$

Assuming h (the number of jurors) = 12, and making n successively 0, 1, 2, 3, 4, 5, the series gives

$$\frac{1}{8192}, \quad \frac{14}{8192}, \quad \frac{92}{8192}, \quad \frac{378}{8192}, \quad \frac{1093}{8192}, \quad \frac{2380}{8192},$$

for the respective probabilities of the error of a verdict when pronounced unanimously by 12 jurors, by a majority of 11 to 1, of 10 to 2, of 9 to 3, of 8 to 4, and of 7 to 5. In the last case the probability of the error is nearly $=\frac{2}{7}$.

85. From these results it appears that the chance of a verdict being wrong which has been pronounced unanimously by twelve jurors is very small; but it is to be remarked, that they have been deduced on the supposition that the unanimity proceeds from agreement in the same opinion, and that the jurors are unbiassed by each other. In this country, where unanimity is compelled by law, the mean probability of a correct verdict can scarcely be considered as greater than that of a verdict pronounced by a simple majority; for, though in most cases the verdict may be supposed to represent the opinion of a larger majority than seven, it may happen, not unfrequently, that a smaller number than five, possessing greater energy or perseverance, may persuade the

others into a surrender of their judgment. In fact, unless the presumption of the guilt of the accused be very great, it would scarcely be possible, without concert, to procure an unanimous verdict in any case. It is also to be observed, that the assumption of all values of u from $\frac{1}{2}$ to 1 being equally probable, may lead to results widely different from the truth. The mean value of u , which depends on the general intelligence of the class of persons from amongst whom the lists of jurors are made up, can only be rightly determined from data furnished by experience. One of the elements, however, which require to be known for this purpose, is the number of jurors who concur in, and dissent from, the verdict. The forced unanimity of the law renders it impossible to obtain this element from the records of the English courts; but in France and Belgium, where the majority and minority are known and recorded, the same obstacle does not exist, and the "*Comptes Généraux de l'Administration de la Justice Criminelle*," published by the French Government, have enabled Poisson to deduce mean values of u and p for that country, and consequently to obtain the necessary data for one of the most interesting applications of the theory of Probabilities. The general results were as follows: During the six years from 1825 to 1830 inclusive, the system of criminal legislation in France underwent no change; the jury consisted of 12, and a simple majority was only required to concur, though when it happened that the majority was the least possible, the Court had power to overrule the verdict. On comparing, according to the rules of the theory, the verdicts given in the cases tried before the criminal courts during those six years, it was found that for the whole of France, the probability (u) of a juror giving a correct verdict was a little greater than $\frac{2}{3}$ with respect to crimes

against the person, and nearly equal to $\frac{1}{17}$ with respect to crimes against property; without distinction of the species of crime, it was found to be a very little below $\frac{3}{4}$. The other element, the probability (p) of the guilt of the accused before the trial, was found not much to exceed $\frac{1}{2}$ (being between 0.53 and 0.54) with respect to crimes against the person, while it a little exceeded $\frac{2}{3}$ in respect of crimes against property. Without distinction of crime, its value was very nearly 0.64.

86. On substituting these values of u and p (namely $u=\frac{3}{4}$, $p=.64$, whence $w=\frac{1}{4}$, $q=.36$) in the formula in (81), and making $m=7$, $n=5$, and consequently

$$U = \frac{1.2.3.....12}{1.2.3...7 \times 1.2.3...5} = 792, \text{ we have } G = \frac{792}{4^{12}}$$

$(3^7 \times .64 + 3^5 \times .36) = .07$ nearly. Hence it may be expected, that in a hundred trials it will happen only seven times that the accused will be pronounced guilty by the smallest possible majority. If $m=12$, and $n=0$, we shall have $u^m p + w^m q = .02027 = \frac{1}{50}$ nearly, for the probability of an unanimous verdict of guilty, and $u^m q + w^m p = .0114$ for the probability of an unanimous verdict of not guilty.

Making the same substitutions in the formula in (82), we have for the probability of a verdict guilty being correct, from which 5 jurors out of 12 dissent, $\varpi_1 = \frac{1}{17}$; and $\varpi_2 = \frac{1}{17}$ for the probability of its being wrong.

Substituting the same values in the series represented by W and W' in (83), and supposing n to have all values from $n=0$ to $n=5$, there results $W = \frac{3^7}{4^{12}} \times 7254$, $W' = \frac{1}{4^{12}} \times 239122$, whence $\frac{Wp}{Wp + Wq} = \frac{126915984}{127992033} = \frac{118}{119}$ nearly. This is the probability that a verdict guilty, pro-

nounced by a majority of seven against five *at least*, is correct. The probability of the same verdict being wrong, is therefore $\frac{1}{119}$; so that out of 119 verdicts, respecting which we know nothing else, than that seven at least of the jury concurred in finding the accused guilty, we may expect one to be wrong, or that one person out of 119 so condemned will be innocent.

SECTION VIII.

OF THE SOLUTION OF QUESTIONS INVOLVING LARGE
NUMBERS.

87. The probabilities of the different compound events which can result from the combination of any number of simple events, $E_1, E_2, E_3, \&c.$ being (13) measured respectively by the several terms of the developement of the multinomial $(p+q+r+\&c.)^n$, the most probable of those compound events will be that which corresponds to the term having the greatest numerical value. Let us consider the case of two simple contrary events E and F , the probabilities of which are respectively p and q , and suppose the number of occurrences to be 4. Neglecting the order of occurrence, the different combinations, with their respective probabilities, are the following :

$$\begin{array}{ccccccc} EEEE, & EEEF, & EEEF, & EFFF, & FFFF, \\ p^4, & 4p^3q, & 6p^2q^2, & 4pq^3, & q^4. \end{array}$$

Now it is evident that the numerical values of these probabilities depend on the ratio of p to q , as well as on the coefficient by which they are multiplied, and that values may be given to p and q , such that any one of the terms may be made the greatest or the least in the series. If we suppose $p=q$, and consequently $p=\frac{1}{2}, q=\frac{1}{2}$, (since $p+q=1$) the probabilities of the different cases become respectively

$$1, \frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16};$$

whence it appears, that the most probable combination is that which corresponds to $6p^2q^2$, or in which each of the simple events occurs twice, the probability of this combination being $\frac{6}{16}$, while that of either of the simple events occurring four times in succession is only $\frac{1}{16}$.

When the number of trials is 5, the probabilities of the several cases are respectively

$$p^5, \quad 5p^4q, \quad 10p^3q^2, \quad 10p^2q^3, \quad 5pq^4, \quad q^5,$$

which, when $p=q$, become

$$\frac{1}{32}, \quad \frac{5}{32}, \quad \frac{10}{32}, \quad \frac{10}{32}, \quad \frac{5}{32}, \quad \frac{1}{32},$$

so that there are two different combinations equally probable, namely, that in which E occurs three times and F twice, and that in which E occurs twice and F three times; and of the six possible combinations these two are the most probable, having in their favour a number of chances twice as great as the two cases in which one of the events occurs only once, and the other four times, and ten times greater than the two cases in which either of the simple events occurs in each of the five trials.

From these two instances it may be inferred in general, that when h is an even number, the most probable compound event is that of which the probability is represented by the middle term of the developement of $(p+q)^h$; and that when h is an odd number, there are two compound events equally probable, and more probable than any other, namely, those corresponding to the terms which occupy the middle of the series, supposing in both cases $p=q$. This supposition gives $(p+q)^h = (1+1)^h (\frac{1}{2})^h$; therefore in the case in which h is an even number, the general expression for the greatest term is

$$\frac{h(h-1)(h-2)\dots(h-\frac{1}{2}h+1)}{1.2.3\dots\frac{1}{2}h} \left(\frac{1}{2}\right)^h;$$

and when h is odd, the general expression for either of the two equal terms, which are greater than any of the other terms, is

$$\frac{h(h-1)(h-2)\dots\{h-\frac{1}{2}(h+1)+1\}}{1.2.3\dots\frac{1}{2}(h+1)}\left(\frac{1}{2}\right)^h.$$

88. When p and q are unequal, the greatest term of the expansion of $(p+q)^h$ will not occupy the middle of the series, but its place may be found by comparing two consecutive terms. Let $h=m+n$. The general term of the series then becomes

$$\frac{1.2.3\dots h}{1.2.3\dots m \times 1.2.3\dots n} p^m q^n;$$

and the term immediately preceding is

$$\frac{1.2.3\dots h}{1.2.3\dots(m+1) \times 1.2.3\dots(n-1)} p^{m+1} q^{n-1}.$$

Dividing the first of these by the second, we get for the quotient $(m+1)q \div np$, which, therefore, is the ratio of two consecutive terms taken at any part of the series. If this ratio be greater than 1, the term which has been taken as the dividend is greater than the preceding one which has been taken as the divisor; and it is evident that the terms must go on increasing, from the beginning of the series, so long as the ratio in question is greater than 1. But if the ratio be less than 1, the preceding term is greater than the succeeding, and the terms will become less and less as they are nearer the end of the series. Let $(m+1)q \div np = 1$; then, since $p+q=1$, and $m+n=h$, we have $n=(h+1)q$, and consequently the ratio of any term to the next preceding is greater or less than 1 according as n is less or greater than $(h+1)q$. Now n is necessarily a whole number; therefore if $(h+1)q$ be a whole number, take $n=(h+1)q$, and the two terms of the series given by the expansion

of $(p+q)^h$, in which the exponents of q are $n-1$ and n , will be equal to each other, and each greater than any other term of the series. But if $(h+1)q$ be not a whole number, let $(h+1)q-x$ be the nearest whole number *less* than $(h+1)q$, and make $n=(h+1)q-x$; then the greatest term of the developement will be that in which the exponent of q is n .

Since $n=(h+1)q-x$, we have $q=(n+x)\div(h+1)$, whence
$$p=1-\frac{n+x}{h+1}=\frac{m+1-x}{h+1}, \text{ and therefore } \frac{q}{p}=\frac{n+x}{m+1-x}.$$

Now x is by hypothesis less than 1, therefore if m and n are large numbers, we have, very nearly, $q:p=n:m$; or, since $m+n=h$, $m=hp$, $n=hq$. It follows therefore, that the greatest term of the developement of the binomial $(p+q)^h$ is that in which the exponents of p and q are to each other in the ratio of p to q , or more nearly in that ratio than are any other two numbers whose sum is h . In other words, the most probable combination of two simple events, E and F, in any number of trials, is that in which the number of occurrences of E is to the number of occurrences of F in the ratio of their respective probabilities.

89. In the same manner it may be shewn, that when there are more than two simple events, of which one must occur in every trial, the most probable result of any number of trials is that combination in which the number of repetitions of each simple event is in proportion to its probability in a single trial. Thus, the probabilities of the simple events being respectively p, q, r , &c. the most probable compound event is that whose probability is expressed by that term of the expansion of $(p+q+r+\&c.)^h$, which has for its argument p^{hp}, q^{hq}, r^{hr} , &c.

90. Having determined the form of the greatest term of

the series, we have next to find a method of approximating to its numerical value ; for its coefficient containing the product of the natural numbers from 1 to h inclusive, its direct calculation becomes impracticable even when h is only a moderately large number. The theorem which gives the approximate value of this product is known by the name of *Stirling's Theorem*, having been discovered by that mathematician. As its investigation is a matter of pure analysis, we shall not stop to give it here, but refer the reader to the *Treatise on Differences and Series*, by Sir John Herschel, in the translation of Lacroix's *Elementary Treatise on the Differential and Integral Calculus*, p. 568.¹ The theorem is as follows : Let x be any number, then

$$1 \cdot 2 \cdot 3 \dots x = x^x e^{-x} \sqrt{2\pi x} \left(1 + \frac{1}{12x} + \frac{1}{288x^2} + \&c.\right)$$

where e is the number of which the Napierean logarithm is unit, or the number 2.71828, and π the ratio of the circumference of a circle to the diameter, or 3.14159.

When x is a large number, the term divided by $12x$ becomes very small, and the series within the brackets may be considered as equal to unity. In this case, then, the formula becomes

$$1 \cdot 2 \cdot 3 \dots x = x^x e^{-x} \sqrt{2\pi x},$$

which gives a sufficient approximation in most cases. If, for example, $x=1000$, the result will be within a 12000th part of the truth.

Now, let E and F be two events of such a nature that the one or the other must happen in every trial ; let p and q be their respective probabilities, and P the probability that in

¹ Stirling's investigation of the theorem, or rather of its equivalent, to find the sum of the logarithms of a series of numbers in arithmetical progression, is given in his *Methodus Differentialis*, p. 135.

$m+n=h$ trials, E will happen m times and F n times ; then by (12) we have

$$P = \frac{1 \cdot 2 \cdot 3 \dots h}{1 \cdot 2 \cdot 3 \dots m \times 1 \cdot 2 \cdot 3 \dots n} p^m q^n.$$

When m , n , and h are large numbers, the value of this coefficient may be computed from the above formula, which gives

$$\begin{aligned} 1 \cdot 2 \cdot 3 \dots h &= h^h e^{-h} \sqrt{2\pi h}, \\ 1 \cdot 2 \cdot 3 \dots m &= m^m e^{-m} \sqrt{2\pi m}, \\ 1 \cdot 2 \cdot 3 \dots n &= n^n e^{-n} \sqrt{2\pi n}, \end{aligned}$$

whence

$$P = \frac{h^h e^{-h} \sqrt{h}}{m^m n^n e^{-(m+n)} \sqrt{(2\pi mn)}} p^m q^n = \left(\frac{hp}{m}\right)^m \left(\frac{hq}{n}\right)^n \sqrt{\frac{h}{2\pi mn}}.$$

This expression represents any term of the series $(p+q)^h$.

The greatest term, which corresponds to the most probable result, is (88) that in which m and n are to each other in the ratio of p to q , or when $m=hp$, and $n=hq$. Let the greatest term therefore be denoted by P_o , that is to say, let P_o be the chance of the most probable result of h trials, and we shall have

$$P_o = \sqrt{(h \div 2\pi mn)}, \text{ or } P_o = \sqrt{(1 \div 2\pi hpq)}.$$

This last formula shews that the absolute probability of that combination which has the greatest number of chances in its favour becomes less and less as the number of trials is increased ; for the fraction $1 \div h$, to the square root of which the probability is proportional, diminishes as h is increased.

91. As an example, suppose a shilling to be tossed 100 times in succession. In this case $p=q=\frac{1}{2}$, $hp=50$, $hq=50$, and the most probable result of the trials is 50 times head and 50 times tail. We have then $h=100$, $m=50$, and $\sqrt{(h \div 2\pi mn)} = 1 \div \sqrt{(50\pi)}$ for the measure of the probabi-

lity that the event will happen in this way exactly. On calculation, this is found $=.07979$; whence it appears, that although 50 heads and 50 tails is a more probable result of 100 trials than any other combination which can be named, its absolute probability is measured by a very small fraction. The probability of the contrary event, or that there will not be thrown 50 heads and 50 tails exactly, is $1 - .07979 = .92021$, so that the odds against the event are about 92 to 8, or 23 to 2. Had the number of trials been 1000, the probability of 500 times head and 500 times tail exactly, though more likely to occur than any other combination, would have been found $1 \div \sqrt{(500\pi)}$; that is to say, $\sqrt{10}$ times, or rather more than 3 times less than in the former case. In general, when the chances in favour of the simple events are equal, the probability of the combination which is more likely to happen than any other, is inversely proportional to the square root of the number of trials.

92. The formulæ in (90) enable us also to determine the ratio of the greatest term of the developement of $(p+q)^h$ to any other term of the series, and consequently the relation of the probabilities of the different compound events. Let $m:n :: p:q$, whence $m=hp$ and $n=hq$, and let P_x denote the probability that in h trials the event E will occur $(m-x)$ times, and the event F $(n+x)$ times, the probabilities of the simple events E and F being respectively p and q . By (13) we have

$$P_x = \frac{1 \cdot 2 \cdot 3 \cdot \dots \cdot h}{1 \cdot 2 \cdot 3 \cdot \dots \cdot (m-x) \times 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n+x)} p^{m-x} q^{n+x},$$

which by (90) becomes

$$P_x = \frac{h^h e^{-h} \sqrt{(2\pi h)}}{(m-x)^{m-x} e^{-(m-x)} \sqrt{2\pi(m-x)} \times (n+x)^{n+x} e^{-(n+x)} \sqrt{2\pi(n+x)}} \\ \times p^{m-x} q^{n+x}; \text{ whence, substituting } m \div h \text{ for } p, \text{ and } n \div h$$

for q , and leaving out the factors common to the numerator and denominator, we find,

$$P_x = \sqrt{\left(\frac{h}{2\pi}\right)} (m-x)^{-m+x-\frac{1}{2}} (n+x)^{-n-x-\frac{1}{2}} m^{m-x} n^{n+x}.$$

Now $\log (m-x)^{-m+x-\frac{1}{2}} = (-m+x-\frac{1}{2}) \log (m-x)$; and

$$\log (m-x) = \log m - \frac{x}{m} - \frac{x^2}{2m^2} - \&c.$$

therefore $\log (m-x)^{-m+x-\frac{1}{2}} =$

$$(-m+x-\frac{1}{2}) \log m - (-m+x-\frac{1}{2}) \left(\frac{x}{m} + \frac{x^2}{2m^2} + \&c. \right);$$

whence, neglecting terms divided by m^2 , m^3 , &c., m being supposed to be a large number in comparison with x ,

$$\log (m-x)^{-m+x-\frac{1}{2}} = (-m+x-\frac{1}{2}) \log m + x - \frac{x^2}{2m} + \frac{x}{2m};$$

therefore, on passing to numbers,

$$(m-x)^{-m+x-\frac{1}{2}} = m^{-m+x-\frac{1}{2}} \times e^{x-\frac{x^2}{2m}} \times e^{\frac{x}{2m}};$$

or, since $e^{\frac{x}{2m}} = 1 + \frac{x}{2m} + \frac{x^2}{2 \cdot 4m^2} + \&c.$

$$(m-x)^{-m+x-\frac{1}{2}} = m^{-m+x-\frac{1}{2}} e^{x-\frac{x^2}{2m}} \left(1 + \frac{x}{2m} \dots \right).$$

In like manner, by changing m into n , and x into $-x$, we get

$$(n+x)^{-n-x-\frac{1}{2}} = n^{-n-x-\frac{1}{2}} e^{-x-\frac{x^2}{2n}} \left(1 - \frac{x}{2n} \dots \right).$$

Multiplying the first of these two expressions by m^{m-x} , and the second by n^{n+x} , we have

$$(m-x)^{-m+x-\frac{1}{2}} \times m^{m-x} = m^{-\frac{1}{2}} e^{x-\frac{x^2}{2m}} \left(1 + \frac{x}{2m} \dots \right)$$

$$(n+x)^{-n-x-\frac{1}{2}} \times n^{n+x} = n^{-\frac{1}{2}} e^{-x-\frac{x^2}{2n}} \left(1 - \frac{x}{2n} \dots \right),$$

whence, substituting these values in that of P_x , and neglecting the quantity divided by mn ,

$$P_x = \sqrt{\left(\frac{h}{2\pi mn}\right)} e^{-\frac{x^2}{2m} - \frac{x^2}{2n}} = \sqrt{\left(\frac{h}{2\pi mn}\right)} e^{-\frac{hx^2}{2mn}}.$$

The term of the series $(p+q)^h$ which corresponds to this value of P_x is that which is x places to the right of the greatest term; and it has been shewn, (90), that the greatest term has for its expression $\sqrt{(h \div 2\pi mn)}$; therefore the greatest term being denoted, as before, by P_o , and the term which comes after it x places by P_x , we have

$$P_x = P_o e^{-hx^2 \div 2mn},$$

that is to say, the probability the event E will happen m times and fail n times in $m + n$ trials, is to the probability of its happening $(m-x)$ times and failing $(n+x)$ times in the ratio of 1 to $e^{-hx^2 \div 2mn}$.

Since the numbers m and n enter symmetrically into the expression $e^{-hx^2 \div 2mn}$, it is evident that the result would have been the same if, instead of seeking the ratio of the greatest term to that which succeeds it by x places, we had sought the ratio of the greatest term to that which precedes it by x places. Hence if the most probable result of $m+n$ trials be that E will happen m times and fail n times, the probability that it will happen $m-x$ times and fail $n+x$ times is the same as the probability that it will happen $m+x$ times and fail $n-x$ times.

The following example will suffice to shew the application of the formula: A die is thrown 6000 times, required the probability that the number of aces turned up will be exactly 960?

Here p , the chance of throwing ace, is $\frac{1}{6}$, $q = \frac{5}{6}$, and $h=6000$; whence $m=hq=1000$, and $n=hq=5000$. We have first to find P_o , the chance of the most probable result,

or of 1000 aces. By (90), $P_o = \sqrt{h \div 2\pi mn}$; whence, substituting the above values, $P_o = \sqrt{3} \cdot \sqrt{(5000 \times 3.14159)}$. On performing the operation indicated by the logarithmic tables, we get $\log P_o = 8.14050$, whence $P_o = .0138$.

The calculation of $e^{-hx^2 \div 2mn}$ is as follows: Assume $t^2 = hx^2 \div 2mn$. We have $x = 1000 - 960 = 40$.

$$\begin{array}{r}
 \log 40 = 1.60206 \\
 2 \\
 \hline
 3.20412 \\
 \log h = \log 6000 = 3.77815 \\
 \hline
 \log hx^2 = 6.98227 \\
 \log 2mn = \log 10,000,000 = 7 \\
 \hline
 \log t^2 = 9.98227 \\
 \log e = .43429, \log .43429 = 9.63778 \\
 \hline
 \log (t^2 \times .43429) = 9.62005 \\
 t^2 \times .43429 = t^2 \log e = .41692 \\
 -t^2 \log e = 9.58308 \\
 \text{add } \log P_o = 8.14050 \\
 \hline
 \log P_x = 7.72358
 \end{array}$$

therefore $P_x = .0053$, which is the chance of 960 aces exactly. The odds against this event are therefore 9947 to 53, or nearly 188 to 1.

93. When h , m , and n are large numbers, and x is small, the exponential $e^{-hx^2 \div 2mn}$ differs little from unity, and it decreases slowly as x increases, so long as x is small in comparison of m and n . Suppose $m = n$ and $x = \sqrt{m}$, it becomes

$$e^{-1} = \frac{1}{e} = \frac{1}{2.7182818}; \text{ so that if we assume } m = 100, \text{ the}$$

10th term before or after the greatest would still exceed the 3d part of the greatest. But when x becomes greater than

\sqrt{m} or \sqrt{n} the exponential, and consequently also the terms which are multiplied by it, begin to diminish with great rapidity, and the diminution is more rapid as x increases. If $m=n=100$, and $x=50$, then the exponent $hx^2 \div 2mn=25$, so that $e^{-hx^2 \div 2mn} = 1 \div e^{25}$, a quantity which is altogether insensible. We may therefore conclude generally that when h is a large number, the principal terms of the development of $(p+q)^h$ are those which are near the greatest term, and that h may be taken so large that the terms towards the beginning or end of the series may at length become smaller than any assignable quantity.

94. From the proposition which has now been demonstrated it follows, that although the probability of that particular compound event which has the greatest number of chances in its favour is very small when the number of trials is great, yet on account of the rapid diminution of the terms towards the beginning and end of the series, the sum of a comparatively small number of terms taken on both sides of the greatest, may be very much greater than all the remaining terms of the series; and, consequently, there will be a very great probability that the compound event will be represented by one or other of those terms. This consideration leads us to one of the most important questions in the theory, namely, to determine the probability that in a large number of trials, h , an event E , which must either happen or fail in each trial, and of which the chance of happening in any trial is p , will happen not less than $hp-l$ times, and not oftener than $hp+l$ times; or, making $hp=m$, $hq=n$, to determine the probability that the number of occurrences of E will be included between the limits $m \pm l$.

Let x be any number between 0 and l . Then (92) the probability that E will occur $(m-x)$ times and fail $(n+x)$

times is $P_x = P_o e^{-hx^2 \div 2mn}$ (where $P_o = \sqrt{h \div 2\pi mn}$). Now if in this expression we make x successively equal to each of the numbers $0, 1, 2, \dots, l$, we shall have the respective probabilities of E happening $m, m-1, m-2, \dots, m-l$ times in h trials; and the sum of these probabilities will be the probability that E happens *not oftener* than m times, and *not seldomer* than $m-l$ times. The same suppositions with respect to x will give the probabilities of E happening $m, m+1, m+2, \dots, m+l$ times, the sum of which will be the probability that E happens *not seldomer* than m times, and *not oftener* than $m+l$ times. Adding, therefore, those two sums, and deducting P_o the probability which corresponds to $x=0$, on account of its being included in each sum, and therefore having been counted twice, the result will be the sum of the terms of the binomial $(p+q)^h$ comprised between, and including, the two terms of which the first has for a factor p^{m+l} , and the last p^{m-l} , and will therefore express the probability that the number of occurrences of E will fall within the limits $m \pm l$. Let this probability be denoted by R, and let SP_x represent the sum of all the values of P_x obtained by substituting successively $0, 1, 2, 3, \dots, l$ for x , we then have $R = 2SP_x - P_o$, whence, writing for P_x and P_o , their values,

$$R = 2S \sqrt{\left(\frac{h}{2\pi mn}\right)} e^{-\frac{hx^2}{2mn}} - \sqrt{\frac{h}{2\pi mn}}.$$

95. In order to find an approximate value of this expression we must have recourse to a formula first given by Euler for converting sums of the kind denoted by S into definite integrals (for which see Lacroix, *Traité du Calcul Différentiel et Intégral*, tom. iii. p. 136, or Herschel's *Treatise on Differences*, p. 513). Assuming u to denote a function of x , the formula is as follows:

$$Su = \int u dx + \frac{1}{2}u + \frac{1}{1 \cdot 2 \cdot 6} \frac{du}{dx} + \&c. + \text{constant}.$$

On making $u = P_x = P_0 e^{-hx^2 \div 2mn}$, we find $\frac{du}{dx} = -\frac{P_0 hx}{mn} e^{-hx^2 \div 2mn}$; therefore, if we suppose x to be not greater than \sqrt{m} or \sqrt{n} , this differential coefficient is of the order $1 \div h$, (as may be easily shewn by substituting hp for m , and hq for n), and may be rejected, since h is supposed to be a very large number. The above equation therefore becomes

$$SP_x = P_0 \int e^{-hx^2 \div 2mn} dx + \frac{1}{2} P_0 e^{-hx^2 \div 2mn} + \text{constant};$$

and on supposing $x=0$ this gives $0 = -\frac{1}{2} P_0 + \text{constant}$, therefore the constant is equal to $\frac{1}{2} P_0$, and we have

$$SP_x = P_0 \int e^{-hx^2 \div 2mn} dx + \frac{1}{2} P_0 e^{-hx^2 \div 2mn} + \frac{1}{2} P_0.$$

Assume $t = x\sqrt{(h \div 2mn)}$, whence $dt = dx\sqrt{(h \div 2mn)}$; substitute these in the above equation, and it becomes by reason of $P_0 = \sqrt{(h \div 2\pi mn)}$,

$$SP_x = \frac{1}{\sqrt{\pi}} \int e^{-t^2} dt + \frac{1}{2} P_0 e^{-t^2} + \frac{1}{2} P_0,$$

whence, from the equation $R = 2SP_x - P_0$, we obtain

$$R = \frac{2}{\sqrt{\pi}} \int e^{-t^2} dt + P_0 e^{-t^2}.$$

The integral in this expression must be taken between the limits $x=0$ and $x=l$. When $x=0$ we have also $t=0$, and when $x=l$, then $t=l\sqrt{(h \div 2mn)}$; therefore assuming τ for the limiting value of t , that is, making $\tau = l\sqrt{(h \div 2mn)}$, and therefore $l = \tau\sqrt{(2mn \div h)}$, there results

$$R = \frac{2}{\sqrt{\pi}} \int_0^\tau e^{-t^2} dt + P_0 e^{-\tau^2}$$

for the probability that the number of occurrences of E will fall between $m \pm \tau\sqrt{(2mn \div h)}$; or, replacing m and n by

hp and hq , the probability that the number of occurrences of E will fall within the limits $hp \pm \tau\sqrt{(2hpq)}$, or be equal to one of those limits.

This expression for R has been found by neglecting quantities of the order of smallness $1 \div h$; consequently the greater the value of h the more nearly it approaches to accuracy, but it is only rigorously true when h is infinite.

96. The integral $\int e^{-t^2} dt$ is computed as follows. Developing the exponential e^{-t^2} in a series of the ascending powers of t^2 , and integrating the successive terms between the limits $t=0$ and $t=\tau$, we find

$$\int_0^\tau e^{-t^2} dt = \tau - \frac{\tau^3}{1 \cdot 3} + \frac{\tau^5}{1 \cdot 2 \cdot 5} - \frac{\tau^7}{1 \cdot 2 \cdot 3 \cdot 7} + \&c.;$$

a series which converges rapidly when τ is less than unity.

In the contrary case, however, or when τ is greater than unity, the series is divergent, and it is necessary to proceed by a different method. Let the factor e^{-t^2} be multiplied and divided by t ; we have then

$$\int e^{-t^2} dt = \int \frac{1}{t} e^{-t^2} t dt,$$

and on integrating by parts

$$\int \frac{1}{t} e^{-t^2} t dt = -\frac{e^{-t^2}}{2t} - \frac{1}{2} \int \frac{1}{t^2} e^{-t^2} dt.$$

Repeating the same process on the last integral, and so on with the last after each succeeding integration, the following series is obtained,

$$\int e^{-t^2} dt = -\frac{e^{-t^2}}{2t} \left\{ 1 - \frac{1}{2t^2} + \frac{3}{2^2 t^4} - \frac{3 \cdot 5}{2^3 t^6} + \&c \right\}.$$

When $t = \infty$ the right hand side of this equation becomes 0; whence between the limits $t=\tau$ and $t=\infty$, we have

$$\int_{\tau}^{\infty} e^{-t^2} dt = -\frac{e^{-\tau^2}}{2\tau} \left\{ 1 - \frac{1}{2\tau^2} + \frac{3}{2^2\tau^4} - \frac{3 \cdot 5}{2^3\tau^6} + \&c \right\},$$

a series which converges very rapidly when τ is greater than unity. Now the value of a definite integral between 0 and infinity is obviously equal to its two parts, of which the first is taken between 0 and τ , and the second between τ and infinity; that is to say,

$$\int_0^{\tau} e^{-t^2} dt = \int_0^{\infty} e^{-t^2} dt - \int_{\tau}^{\infty} e^{-t^2} dt.$$

But $\int_0^{\infty} e^{-t^2} dt$ is well known¹ to have for its expression $\frac{1}{2} \sqrt{\pi}$, therefore

$$\int_0^{\tau} e^{-t^2} dt = \frac{1}{2} \sqrt{\pi} - \int_{\tau}^{\infty} e^{-t^2} dt,$$

so that the integral may be computed from either of the above series, according as τ is less or greater than 1.

The integral $\int e^{-t^2} dt$ is of great importance in the higher mathematics. It occurs in the investigation of the path of a ray of light through the atmosphere, and of the law of the diffusion of heat in the interior of solid bodies, as well as in the determination of the degree of reliance that may be placed on the results of astronomical observations, and generally in most of the more difficult and important applications of the theory of probabilities. A table of its values from $t=0$ to $t=3$, for intervals each $=.01$, was given by Kramp, at the end of his *Analyse des Refractions Astro-*

¹ The formula $\int_0^{\infty} e^{-t^2} dt = \frac{1}{2} \sqrt{\pi}$ is ascribed by Gauss (*Theoria Motus Corporum Caelestium*, p. 212,) to Laplace. On making $e^{-t^2} = z$, the integral is transformed into $\frac{1}{2} \int dz \left(\log \frac{1}{z} \right)^{-\frac{1}{2}}$ between the limits $z=0$ and $z=1$, the value of which $=\frac{1}{2} \sqrt{\pi}$, had been given by Euler, long before, in the *Petersburg Memoirs*. See Legendre, *Exercices du Calcul Intégral*, tom. i. p. 301.

nomiques, Strasburg, 1799. In the *Berliner Astronomisches Jahrbuch* for 1834, there is also a table of its values from $t=0$ to $t=2$ (for the same intervals) multiplied by $2 \div \sqrt{\pi}$, with their first and second differences, for the purpose of facilitating interpolation. This last table, which appears to have been derived from that of Kramp, and which is immediately applicable in the calculation of the probability R, we have extended to $t=3$, and given at the end of the present article. As the function which is thus tabulated will occur frequently in what follows, we shall in future, for convenience in printing, denote it by Θ , that is to say, we shall assume

$$\Theta = \frac{2}{\sqrt{\pi}} \int_0^{\tau} e^{-t^2} dt = 1 - \frac{2}{\sqrt{\pi}} \int_{\tau}^{\infty} e^{-t^2} dt,$$

the two forms being equivalent in consequence of the above equation.

97. Some very important conclusions follow immediately from the formula in (95). The quantity R denotes the probability that in a very great number of trials h , the event E, of which the *a priori* probability in any trial is p , will occur not seldomer than $hp-l$ times, and not oftener than $hp+l$ times, or that the number of its occurrences will be included between the limits $hp \pm l$, or at least be equal to one of those limits. Hence R also denotes the probability that the ratio of the occurrences of E to the whole number of trials, will be included within the limits $p \pm l \div h$. We have assumed $\tau = l \sqrt{(h \div 2mn)}$; but $m = hp$ and $n = hq$; therefore $\tau = l \div \sqrt{(2hpq)}$, whence $l = \tau \sqrt{(2hpq)}$, and consequently $l \div h = \tau \sqrt{(2pq \div h)}$. Now, if we suppose τ to be constant, so that the probability expressed by R may remain the same, then p and q being given, l is proportional to

the square root of h , and consequently the greater the number of trials the smaller will l be in proportion to that number. Thus, if the number of trials be 1000, and we have a given probability R that the number of occurrences of E will not differ more than 10 from the number which is the most probable of all (that is, from $1000p$), then if we take 100,000 trials, we shall have the same probability R that the number of occurrences of E will not differ more than $10 \times \sqrt{100} = 100$ from the most probable number. But a difference of 10 in 1000 is 1-100th of the whole, whilst a difference of 100 in 100,000 is 1-1000th of the whole, and thus the ratio of l to h becomes smaller and smaller, or the ratio of the occurrences of E to the whole number of trials approaches nearer and nearer to p , as the number of trials is increased; and the experiments may be repeated until the difference between p and $p \pm l \div h$, in respect of a given probability R which may be as great as we please, shall be less than any assignable quantity.

If, on the other hand, we suppose $l \div h$ to be constant, then τ is proportional to the square root of the number of trials. But as τ increases, Θ , and consequently R , approaches nearer and nearer to unity, (and it may be seen, by referring to the table, that it is only necessary to have $\tau = 3$ in order to have $\Theta = .9999779$); whence the number of trials h may always be increased until we obtain a probability approaching as nearly to certainty as we please, that the number of occurrences of E will be comprised within the given limits ($hp \pm l$); or, which is the same thing, that the ratio of the number of occurrences of E to the whole number of trials, shall not differ from p , the probability of E in a single trial, more than a given quantity $l \div h$ which may be less than any assigned fraction. This is the celebrated theorem

which was demonstrated by James Bernoulli in the *Ars Conjectandi*.

98. The application of the preceding results to numerical examples, is rendered extremely easy by means of the table of the values of Θ . From the formula in (95) we have the probability

$$R = \Theta + P_0 e^{-\tau^2},$$

that the occurrences of E in h trials will fall within the limits $hp \pm l$, the relation between l and τ being given by the equation $l = \tau \sqrt{2hpq}$. If, therefore, we suppose l to be given, τ becomes known, and the corresponding value of Θ is found from the table; and, conversely, if Θ be assumed, τ is given by the table, whence the corresponding limits l are deduced. With respect to the quantity $P_0 e^{-\tau^2}$, we may observe that it denotes the probability that the number of occurrences of the event E will be $hp + l$ or $hp - l$ precisely (92), and is therefore always a very small fraction when h is a large number (90). It may be regarded as a correction of Θ , which in most cases might be omitted without sensibly affecting the result; but when h is not very large, or l is a small number, it becomes necessary to take it into account. In such cases its value may be computed directly as in the example in (92); but this labour may be avoided by increasing τ , so as to include it within the limits of the integral Θ . Thus, let R be the probability that the number of arrivals of E will be included within the limits $hp \pm x$, and R' the probability of the limits being $hp \pm (x+1)$, and let Θ and Θ' be respectively the corresponding values of the integral. We have then, giving P_x the same signification as in (92), the two equations

$$R = \Theta + P_x, \quad R' = \Theta' + P_{x+1},$$

and the difference $R' - R$ of these two probabilities is obviously the double of the probability that the result of the trials will be either $(hp + x + 1)$ times E, or $(hp - x - 1)$ times E, exactly. But the chance of either of these events being P_{x+1} , we have therefore $R' - R = 2P_{x+1}$. Now, when h is large, P_x and P_{x+1} are very small, and very nearly equal to each other, (their difference is in fact of the order of quantities omitted); hence $R' - R = \Theta' - \Theta$, and also $2P_{x+1} = 2P_x$, and consequently $\Theta' - \Theta = 2P_x$, or $P_x = \frac{1}{2}(\Theta' - \Theta)$. Substituting this value of P_x in the equation $R = \Theta + P_x$, we get $R = \frac{1}{2}(\Theta' + \Theta)$; so that if we take from the table the values of Θ' and Θ corresponding to l and $l + 1$, half their sum will give R . But as the interval between Θ' and Θ in the table is always small, half their sum will not differ sensibly from the value of Θ corresponding to $l + \frac{1}{2}$, whence this value of Θ is equal to R , and we have the following rule for determining the limits corresponding to a given probability, or *vice versa* :—

When the limits are assumed, find τ from the equation $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$; then the value of Θ in the table, corresponding to τ is the probability that in h trials the number of occurrences of the event E, the chance of which in a single trial is p , will lie within the limits $hp \pm l$ both inclusive. Conversely, when Θ is assumed, find the corresponding value of τ in the table, by means of which the limit l will be given by the equation $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$. It is obvious, that if the limit l and the probability Θ be both assumed, then h may be determined from the same equation.

99. We will now give some examples of the application of the preceding formulæ.

Suppose $p = q = \frac{1}{2}$, and $h = 200$, and let it be proposed to assign the limits within which there is a probability $= \frac{1}{2}$

that the number of occurrences of E will fall. In this case the equation $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$ becomes $l + \frac{1}{2} = \tau\sqrt{100} = 10\tau$. Now, it is easily found from the table that for $\Theta = \frac{1}{2}$ we have $\tau = .4769$, whence $l + \frac{1}{2} = 4.769$, and $l = 4.269$. On tossing a shilling 200 times, it is therefore more than an even wager that head will turn up not seldomer than 95 times, and not oftener than 105 times.

Suppose $p = q = \frac{1}{2}$, $h = 3600$, and let it be proposed to assign the probability that the number of occurrences of E will not exceed the limits 1800 ± 30 . In this case the equation $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$ becomes $30.5 = \tau\sqrt{(2 \times 900)} = 30\tau\sqrt{2}$, whence $\tau = 30.5 \div 30\sqrt{2} = .7189$; and the table gives $\Theta = .6907 = \frac{28}{41}$ nearly. Hence in tossing a shilling 3600 times, the odds are 28 to 13 that head will not turn up oftener than $1800 + 30 = 1830$ times, nor seldomer than $1800 - 30 = 1770$ times. Neglecting the second term of R (95) and taking simply $l = 10\tau$, the table gives $\Theta = .6827$, which is the solution given by Demoivre, p. 245.

Suppose $p = \frac{1}{6}$, $q = \frac{5}{6}$, and let it be proposed to determine how many trials must be made in order that it may be one to one that the number of occurrences of E will not differ more than 10 from the most probable number.

For $\Theta = \frac{1}{2}$ we have $\tau = .4769$; therefore the equation $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$ becomes $10.5 = .4769\sqrt{(10h \div 36)}$, whence $h = 3.6(10.5 \div .4769)^2$. On computing this formula h is found $= 1745.2$. Say 1746, $\frac{1}{6}$ of which is 291; and it follows that if a die be thrown 1746 times it is an even wager that the number of aces will fall between 291 ± 10 , that is, between 281 and 301, or be equal to one of those numbers.

In (92) we found the probability to be .0053, that in 6000 throws of a die the number of aces will be exactly

960. Let it now be proposed to assign the probability Θ , that in 6000 throws the number of aces will lie between 960 and 1040, that is, between 1000 ± 40 . Here $h=6000$, $p=\frac{1}{6}$, $q=\frac{5}{6}$, and $l=40$; the equation of the limits therefore becomes $40.5 = \tau\sqrt{(10000 \div 6)}$, whence $\tau = .405\sqrt{6} = .992$, corresponding to which the table gives $\Theta = .8394$.

The following question is discussed by Nicolas Bernoulli in the Appendix to Montmort's *Analyse des Jeux de Hazard*, and is noticed by Demoivre and Laplace. From the observations of the births of both sexes in London during 82 years (from 1629 to 1711) it was found that the average number of children annually born in London, was about 14,000, and the ratio of the number of males to that of females, was nearly as 18 to 17, the average number of male births being 7200, and of female births 6800. In the year in which the greatest difference from this ratio took place, the actual numbers were 7037 males and 6963 females, so that the difference from the average amounted to 163. Assuming, then, the comparative facility of male and female births to be as 18 to 17, required the probability that out of 14000 children born, the number of males shall not be greater than 7363, nor less than 7037.

This question is evidently equivalent to the following :—Let 14000 dice, each having 35 faces, 18 white and 17 black, be thrown; what is the probability that the number of white faces turned up, will be comprised between the limits 7200 ± 163 . We have therefore $h=1400$, $p=\frac{18}{35}$, $q=\frac{17}{35}$, $l=163$, and the formula $l + \frac{1}{2} = \tau\sqrt{(2h p q)}$ becomes $163.5 = \tau\sqrt{(2 \times 14000 \times 18 \times 17) \div 35}$, whence $\tau = 1.955$. The corresponding value of Θ is found from the table = .9943, which is the probability that the number of white faces shall

not be greater than 7363, nor less than 7037. The odds in favour of the event are therefore 9943 to 57, or about 175 to 1.

100. We now proceed to consider the case in which the probabilities of the simple events are not known, *a priori*, but inferred from the results of experience. It was shewn in (52) that the probability Π of an event happening m' times, and failing n' times in h' trials, ($h' = m' + n'$), when it has been observed to happen m times, and fail n times in h previous trials, is expressed by this equation

$$\Pi = U \frac{[m+m'] [n+n'] [h+1]}{[m] [n] [h+h'+1]}.$$

Now, when m, n, m', n' , are large numbers, an approximate value of Π , more accurate in proportion as those numbers become larger, is obtained from Stirling's theorem (90), which for any number x gives $[x] = x^x e^{-x} \sqrt{(2\pi x)}$. Applying the theorem therefore to the expressions within the brackets in the above equation, and assuming

$$K = \frac{h+1}{h+h'+1} \sqrt{\frac{(m+m')(n+n')(h+1)}{mn(h+h'+1)}},$$

we obtain, in consequence of $m+n=h$,

$$\Pi = U' K \frac{(m+m')^{m+m'} (n+n')^{n+n'} (h+1)^h}{m^m n^n (h+h'+1)^{h+h'}}.$$

Let $m' = \theta m$, $n' = \theta n$, and consequently $h' = \theta h$; then taking

$\frac{h^h}{(h+h')^{h+h'}}$ for $\frac{(h+1)^h}{(h+h'+1)^{h+h'}}$ (which may be done without sensible error, since h is by supposition a large number,) the equation becomes

$$\Pi = U' K \frac{m(1+\theta)^{m+m'} n(1+\theta)^{n+n'} h^h}{m^m n^n (1+\theta)^{h+h'}};$$

or, since $m+n=h$, $m'+n'=h'$,

$$\Pi = UK \frac{m^{m'} n^{n'} (1 + \theta)^{h+h'}}{h^h (1 + \theta)^{h+h'}} = UK \left(\frac{m}{h}\right)^{m'} \left(\frac{n}{h}\right)^{n'}.$$

Making the same substitutions in the expression denoted by K, we get, after reduction, $K = 1 \div \sqrt{1 + \theta}$; whence,

$$\Pi = \frac{U'}{\sqrt{1 + \theta}} \left(\frac{m}{h}\right)^{m'} \left(\frac{n}{h}\right)^{n'}.$$

The value of Π now found, is the probability that in a future series of trials the ratio of the occurrences of E to those of F will be the same as in the preceding trials, which are supposed to have been very numerous. If the chances of E and F had been given *a priori* equal to $m \div h$ and $n \div h$ respectively, the probability of m' times E, and n' times F in $m' + n'$ future trials would have been $P = U' \left(\frac{m}{h}\right)^{m'} \left(\frac{n}{h}\right)^{n'}$ by (12); hence (since $m' \div h' = m \div h$ and $n' \div h' = n \div h$), the relation between the probability P_o of that combination of simple events which has the greatest number of chances in its favour, when the chances of the simple events are known *a priori*, and the probability of the same combination when the chances of the simple events are only presumed from previous trials, is expressed by this equation,

$$\Pi = P_o \div \sqrt{1 + \theta}.$$

101. When h' is very small in comparison of h , θ becomes a very small fraction, and may be neglected, and we have then $\Pi = P_o$. But when h' is a number comparable with h , Π is less than P_o ; and it diminishes rapidly when θ exceeds 1. The reason of this is obvious. If the contents of the urn are not known *a priori*, however numerous the trials may have been there is only a presumption that the chance of drawing a white ball in a single trial is measured by $m \div h$; whereas, in the case of the ratio of the balls being

previously known, the measure of the probability is certain. As an instance of the manner in which the probability of an assigned series of future events diminishes, when the probabilities of the simple events are inferred from experience, let us suppose $h'=h$, whence $\theta=1$, and consequently $\Pi=P_o \div \sqrt{2}=.7071 \times P_o$. Now it was shewn in (91) that if a ball be drawn at random 100 times from an urn which contains an equal number of black and white balls, the probability P_o , that the result will be 50 white balls, and 50 black, precisely, is .07979. It follows therefore, that if the contents of the urn be unknown, and we can only judge of the relative numbers of the two sorts of balls it contains from having observed that in 100 trials there have been drawn 50 white balls and 50 black, the probability Π of that combination in 100 future trials, becomes $.07979 \times .7071=.05642$.

102. The result obtained in (100) enables us to determine the probability that the number of occurrences of E in h' future trials, will not differ in excess or defect from the most probable number, by more than a certain given number l . It has been shewn (95) that in the case of the probabilities p and q of the simple events being given *a priori*, if we determine τ from the equation $l=\tau\sqrt{(2hpq)}$, the formula

$$R=\Theta+\sqrt{(1\div 2\pi hpq)}e^{-\tau^2}$$

gives the probability R that m will be comprised within the limits $hp\pm\tau\sqrt{(2hpq)}$; or, dividing by h , the probability that the ratio of m to h will be comprised within the limits $p\pm\tau\sqrt{(2pq\div h)}$. Conversely¹, when p and q are not known,

¹ This inference, though admitted by both Laplace and Poisson, is not strictly correct. In a paper published in the *Transactions of the Cambridge Philosophical Society*, (vol. vi. part iii.) Mr. De Mor-

but the event E has been observed to happen m times in h trials, then

$$R = \Theta + \sqrt{(h \div 2\pi mn)} e^{-\tau^2}$$

gives the probability R that p is comprised within the limits

$$\frac{m}{h} \pm \frac{\tau}{h} \sqrt{\frac{2mn}{h}}.$$

These limits approach more nearly to each other as h increases; and when h is a large number, the ratios $m \div h$, $n \div h$ may be assumed, without sensible error, as the chances of E and F in computing the probable result of a future series of h' trials, provided, however, that h' (though absolutely a large number) be small relatively to h . When this condition is not fulfilled, the assumption of $m \div h$ and $n \div h$ as the *a priori* chances of E and F, might lead to considerable error; but an approximation to the limits corresponding to a given value of R may be obtained from the following considerations:—

Suppose a large number h of events to have been observed, and that the result of the observation gave m times E and n times F. Let a new series of h' trials be made, and suppose that in this new series p is the real chance of E

gan has shewn by a direct analysis that in the case of p and q not being known *a priori*, but made equal to the observed ratios $m \div h$, $n \div h$, the presumption of the true value of p lying within the limits stated in the text is not, as there inferred, $\Theta + \frac{1}{\sqrt{(2\pi hpq)}} e^{-\tau^2}$,

but $\Theta - \frac{13p^2 - 13p + 1}{6hpq\sqrt{\pi}} e^{-\tau^2}$. The last correction to Θ is smaller

than the former, and being divided by h , is of the order of quantities that have been rejected in the approximations. It is right to state that the method of simplifying the calculation of R in the direct case, by taking the integral Θ between limits corresponding to $l + \frac{1}{2}$ instead of l , is noticed, for the first time so far as we are aware, by Mr. De Morgan in the same paper.

and q of F ; we have then a given probability R that the number of occurrences of E will fall within the limits $h'p \pm \tau\sqrt{(2h'pq)}$. Now, for p and q substitute the ratios observed in the first set of experiments, namely, $m \div h$ and $n \div h$, and the limits corresponding to R become

$$\frac{mh'}{h} \pm \frac{\tau}{h} \sqrt{(2h'mn)},$$

which, therefore, are the true limits on the hypothesis that the chance of E in a single trial is $m \div h$. But as this chance is not certain, but only presumed, the limits require to be extended in order that R may preserve the same value. Confining our attention to Θ , the first term of the expression for R (the second may be disregarded in the present approximation), let $h' = m' + n'$ and $m' : n' = m : n$, then Θ is the sum of the terms of the binomial $(p+q)^{h'}$ from that in which the exponent of p is $m' + l$ to that in which the exponent is $m' - l$. Now, when p and q are given *a priori*, the chance of m' times E and n' times F in h' trials is P_o ; and when p and q are only presumed from the results of previous trials, the chance of the same combination is Π ; and $(100)\Pi$ is less than P_o in the ratio of 1 to $\sqrt{(1+\theta)}$. In like manner, the chance of each of the other combinations of E and F included in the integral Θ will be less in the case of p and q presumed, than in the case of p and q given, in the same ratio of 1 to $\sqrt{(1+\theta)}$. But it has been seen (93) that when h' is a large number, the terms of the development of $(p+q)^{h'}$ which are nearest the greatest term, diminish at first very slowly; and, further, that only a small number of terms on each side of the greatest are required to be taken, since l is less than $\sqrt{m'}$ or $\sqrt{n'}$ (95); we may therefore, without sensible error, assume Θ to be proportional to the number of terms included in the summation, or

that the value of Θ will not be changed if we include in the summation a number of terms greater in proportion as the value of each individual term is less. Hence it follows that the limits must be increased in the ratio of $\sqrt{1+\theta}$ to 1, and the value of Θ corresponding to τ will give the probability that the number of events E, in h' future trials, will be included between

$$\frac{mh'}{h} \pm \frac{\tau}{h} \sqrt{(2h'mn(1+\theta))}.$$

* 103. The following question may be proposed as an example of the application of the last formula. Out of a given number h of individuals taken at the age A, it has been observed that m are alive at the age $A+a$; required the probability that out of h' other individuals taken at the same age A the number who survive at the age $A+a$ will be included between $m' \pm l$, the ratio of m' to h' being the same as that of m to h .

To solve this question, we have to find τ from the equation $l = \frac{\tau}{h} \sqrt{(2h'mn(1+\theta))}$; and the corresponding value of Θ in the table, will give the required probability.

From the table given in the article MORTALITY, vol. xv. p. 555, it appears that out of 5642 individuals taken at the age 30, the number surviving at the age 50, according to the Carlisle Table, is 4397. Taking those numbers as an example, we have $h=5642$, $m=4397$, $n=1245$; and assuming also $h'=5642$, whence $\theta=1$ and $\sqrt{1+\theta}=\sqrt{2}$, the equation of the limits becomes $l=\tau \times 62.30$. Let it be proposed to determine l from the condition $\Theta=\frac{1}{2}$. In this case the table gives $\tau=.4769$, and we have consequently $l=29.7$. Hence it appears, that if it has been observed that of 5642 individuals taken at the age of 30, 1245 die

before reaching the age of 50, it is an even wager that out of 5642 other individuals also taken at the age of 30, and subjected to the same chances of mortality, the number who die before reaching the age of 50 will lie between 1245 ± 30 , that is, between 1215 and 1275.

104. The following experiment recorded by Buffon, in his *Arithmetique Morale*, affords an example of the application of the preceding formulæ to the determination of the probable existence of a physical cause from the results of a large number of observations. A piece of money was tossed 4040 times successively, and the result was head 2048 times, and tail 1992 times. Supposing the piece to have been perfectly symmetrical, the most probable result would have been the same number of heads and tails. Let it now be proposed to assign the probability afforded by the experiment that the piece was not symmetrical, and that its form or physical structure was such as to render head an event, *a priori*, more probable than tail.

In this case $h=4040$, $m=2048$, $n=1992$; and by (102) we have the probability R (or Θ , neglecting the correction) that p , the unknown chance of head, is comprised between the limits $\frac{m}{h} \pm \frac{\tau}{h} \sqrt{\frac{2mn}{h}}$. Now $\frac{m}{h} = \frac{2048}{4040}$

$=.50693$, and $\frac{\tau}{h} \sqrt{\frac{2mn}{h}} = \tau \times .011124$; therefore if we assume $\tau \times .011124 = .00693$, we shall have the probability Θ that p is comprised between the limits $.50693 \pm .00693$, that is, between two limits of which the least is .5, or one-half. This assumption gives $\tau = .00693 \div .011124 = .623$; and the corresponding value of Θ is found from the table $=.62170$. Now if p lie between the above limits, its value is evidently greater than $\frac{1}{2}$; but the probability of its lying

between those limits is not the whole probability that p is greater than $\frac{1}{2}$; for there is a chance of its exceeding the greatest limit, in which case also its value will be greater than $\frac{1}{2}$. The probability that p is *not* comprised between the assumed limits is $1 - .62170 = .37830$; and if it is not comprised between these limits, there is an equal chance of its being greater than the greatest limit, or less than the least; the probability of its exceeding the greatest limit is consequently $\frac{1}{2} \times .37830 = .18915$. Hence the whole probability that p is greater than .5, or that the chance of head is greater than that of tail is $.62170 + .18915 = .81085$; and the odds are therefore 81 to 19, or rather more than 4 to 1 that the piece was not perfectly symmetrical.

105. The formulæ which have been demonstrated in the present section are immediately applicable to the determination of the probable limits of the gain or loss which may arise from undertaking a great number of risks with a given expectation in respect of each. The following question has important practical applications. A is interested in a great number of similar enterprises, in each of which E or F must necessarily happen. When E happens he receives the sum a , and when F happens he pays the sum β ; required the probability that his gain or loss shall be comprised within given limits?

Let p be the chance of the event E, q that of F, and h the number of enterprises. Suppose E happens m times, and F n times; the sum to be received will be ma , and the sum to be paid will be $n\beta$, and therefore his gain will be $ma - n\beta$. Let $m = hp$, $n = hq$, then m times E and n times F is the most probable result, and in this case the gain $ma - n\beta$ becomes $h(pa - q\beta)$. Find τ from $l + \frac{1}{2} = \tau\sqrt{(2hpq)}$, then (98) Θ is the probability that the number of occur-

rences of E will lie between the limits $hp \mp l$. But if E happens $hp - l$ times, and consequently F $hq + l$ times, the corresponding benefit is $(hp - l)a - (hq + l)\beta = h(pa - q\beta) - l(a + \beta)$; and if E happens $hp + l$ times, and F $hq - l$ times, the benefit is $(hp + l)a - (hq - l)\beta = h(pa - q\beta) + l(a + \beta)$; whence Θ is the probability that his gain, that is, the difference between what he receives and what he pays, will be included within the limits $h(pa - q\beta) \mp l(a + \beta)$ both inclusive.

106. The following conclusions follow immediately from this solution.

(1). If pa be greater than $q\beta$, so that A has a mathematical advantage (however small) in each risk, the risk may be repeated a sufficient number of times, or h may be taken a sufficiently high number, to give a probability as nearly equal to certainty as we please, that A's gain shall exceed any given sum, however great.

(2). Let there be two players A and B, whose chances of gaining a game are respectively p and q , and let β be the sum staked upon each game by A, and a the sum staked by B, then pa is the mathematical expectation of A in respect of a single game, and $q\beta$ that of B; and if pa be greater than $q\beta$ (however small the difference) the game may be repeated so often as to give rise to a probability approaching as nearly to certainty as we please, that A's gain shall become equal to the whole of B's capital, and, consequently, that B will be ruined.

(3). If the mathematical expectations of the two players be equal, then $pa - q\beta = 0$, and the most probable individual result of a large number of games, is that the gains and losses on either side shall be the same. But if l be supposed constant, then τ is inversely proportional to \sqrt{h} , and consequently the game may be repeated until Θ , the probability that the

gain or loss $l(a+\beta)$ shall be comprised within given limits, shall become as small as we please. Hence $1-\Theta$, the probability that the gain or loss shall *not* be comprised within given limits, may be rendered as great as we please; and it follows that although the play may be on terms of perfect equality, it may be continued until a probability shall be obtained, approaching as nearly to certainty as we please, that one of the two players shall be ruined.

(4). The number of games which must be played, to afford a given amount of probability that one of the parties shall lose the whole of his fortune, depends on the magnitude of the stakes $(a+\beta)$; but whether the stakes be large or small, the final result is the same. When the stakes are small, a greater number of games must be played.

107. As an example of this class of problems, we may take the following question: A and B engage in play with equal chances of winning, and stake five sovereigns on each game; how many games must they undertake to play in order that it may be two to one that one of them shall lose *at least* 100 sovereigns?

Here $p=\frac{1}{2}$, $q=\frac{1}{2}$; $a=5$, $\beta=5$, and $l(a+\beta)=100$, whence $l=10$. The equation $l+\frac{1}{2}=\tau\sqrt{(2hpq)}$ therefore becomes $10\cdot5=\tau\sqrt{(h\div2)}$, whence $h=2\times(10\cdot5)^2\div\tau^2$. Now, the odds being 2 to 1 against the limits of the gain or loss not exceeding 100, the probability Θ of the limits not exceeding 100 is $\frac{1}{3}=0\cdot33333$, corresponding to which the table gives by interpolation $\tau=0\cdot30458$; substituting which in the above equation we find $h=2376\cdot8$; so that if 2377 games are played, the odds are 2 to 1 that one of the players shall have gained at least 10 games more than half that number, and, consequently, that the other shall have gained at least 10 less than half, or that one of them shall have gained at least

20 games more than the other, and consequently have gained at least 100 sovereigns.

It is to be carefully observed that this question supposes the account between A and B not to be balanced until 2377 games have been played. If the condition of the play had been that it should cease as soon as A or B should have lost 100 sovereigns, the question would have been of an entirely different kind, and a much smaller number of games would have given the same probability of an equal loss.

108. The question just alluded to belongs to a class of problems connected with the *Duration of Play*, of extreme difficulty, and which have given rise to some of the most abstruse and refined researches in the modern analysis. In order to give an idea of the subject, we may take the following question, which has been frequently considered.

A and B, whose chances of winning a game are respectively p and q , play on these terms: A has m counters, and B has n counters; when A loses a game he gives a counter to B, and when B loses a game he gives a counter to A, and the play is to cease when one of them has lost all his counters. What is the probability that the play, which may go on for ever, shall be finished before more than k games shall have been played.

To take a simple case, suppose each to have three counters, and let the probability be required that the play shall be concluded with or before the ninth game. As the play cannot end with less than three games, let the binomial $(p+q)^5$ be developed, and the terms

$$p^5 + 3p^2q + 3pq^2 + q^5$$

give the respective probabilities of all the cases which can arise in three games. The first term is the probability of A gaining all the three games, the last term is the proba-

bility of B gaining them, and the sum of the remaining two terms is the probability that neither will win all the games, or the chance that a fourth will be played. Now, if the fourth game be played, p is A's chance of winning it, and q B's chance; but these chances will only exist in respect of the fourth game, provided the play be not concluded with the previous one, the probability of which is $3p^2q + 3pq^2$. Multiplying, therefore, $3p^2q + 3pq^2$ by $p + q$, the product

$$3p^5q + 6p^2q^2 + 3pq^5$$

gives the respective probabilities of the different ways in which the four games may be gained by A and B, excepting the two ways in which the play would have terminated with the third game. But the play cannot end in any of these ways; for, taking the first term for example, if B gains a counter before A gains three, the play cannot terminate until A gain back that counter, and three others besides, so that five games must be played. In fact, it is obvious that there is no way of gaining an odd number of counters in an even number of games, or *vice versa*. The last product therefore expresses the chance of the 5th game being played; and by reason of $p + q = 1$ it is equal to $3p^2q + 3pq^2$, the chance of the 4th being played, as it obviously ought to be, since the play cannot terminate with the 4th. Again, if the 5th game be played, p is A's chance of gaining it, and q B's chance of gaining it; multiplying therefore the last product by $p + q$, the different terms of the result, namely,

$$3p^4q + 9p^3q^2 + 9p^2q^3 + 3pq^4$$

give the respective probabilities of all the cases which can arise by the 5th game. The first term is the probability of A gaining 4 games and B gaining 1, and the last term is the

probability of B gaining 4 and A gaining 1. These terms therefore are the probabilities of the play ending in favour of A and B respectively with the 5th game, and the sum of the other two terms is the probability that the play will not terminate with the 5th game, or the chance of another game being played.

By pursuing the same reasoning it will be evident that on rejecting the two extreme terms of the above product, and multiplying the remainder by $p+q$, there will result the probabilities of the different ways in which six games may be played without the one player gaining all the counters of the other. But as the play cannot terminate with the 6th game, multiply again by $p+q$, and the result

$$9p^5q^2 + 27p^4q^3 + 27p^5q^4 + 9p^2q^5$$

will indicate the probability of the different cases that can arise out of the 7th game. Rejecting the two extreme terms, which give the respective probabilities of the play being concluded in favour of A or B, and multiplying the remaining two first by $p+q$ to obtain the different probabilities in respect of the 8th game, and again by $p+q$, as the play cannot terminate with the 8th, we have the product

$$27p^6q^3 + 81p^5q^4 + 81p^4q^5 + 27p^3q^6,$$

of which the first and last terms give the respective chances of A and B winning at the 9th game, and the sum of the other two terms the probability that the play will not be concluded by the 9th.

If we now collect the terms which have been set aside in the successive products, and denote by a and b the respective probabilities of A and B gaining at the 9th game, *or sooner*, we shall have

$$a = p^3 + 3p^4q + 9p^5q^2 + 27p^6q^3,$$

$$b = q^3 + 3q^4p + 9q^5p^2 + 27q^6p^3,$$

where the law of the series is evident.

It is easy to see that this process may be applied whatever be the number of counters which A and B have at the commencement, and whatever be the number, h , of games to which the play is limited. The general rule is as follows: of the two numbers m and n , let m be that which is not less than the other. Raise $p+q$ to the power n , and reject the first term (which gives the chance of A winning n games in succession), and also the last if $m=n$. Multiply the remainder ($h-n$) times in succession by $(p+q)$, rejecting at each multiplication the first or last term of the product when it gives a combination which would terminate the play in favour of A or B; the sum of the terms rejected from the left-hand side of the different products gives the probability in favour of A, and the sum of the terms rejected from the right-hand side the probability in favour of B. As the coefficients of the successive products are obviously formed by adding the coefficient of the corresponding term in the preceding product to that of the term immediately before it, the products may be written down at once without the trouble of multiplication; but it is evident that when m , n , and h are large numbers, it would be quite impracticable to sum the series formed of the rejected terms by the ordinary methods. From the manner in which the series are derived, they are called *recurring* series; a general theory of which was first given by Demoivre in his *Doctrine of Chances*, and forms the most remarkable portion of that work.

109. The general problem is reduced to an equation of finite differences as follows: Let $y_{x,t}$ represent A's expec-

tation when x games have been played, and he has still t counters to win, or B has t counters in his hand. If A gain the next game the value of his expectation will become $y_{x+1, t-1}$, and the chance of his gaining it is p ; therefore his expectation in respect of that event is $py_{x+1, t-1}$. On the other hand, if A loses the next game his expectation will become $y_{x+1, t+1}$, and the chance of losing it is q ; therefore his expectation in respect of that event is $qy_{x+1, t+1}$. Hence, according to the principles laid down in (32),

$$y_{x, t} = py_{x+1, t-1} + qy_{x+1, t+1}$$

a linear equation of finite differences, with three independent variables. It is therefore on the integration of an equation of this kind that the problem of the duration of play ultimately depends, but the subject is of much too complicated a nature to admit of its being satisfactorily explained in this place. We must therefore content ourselves with referring the reader to the treatise on generating functions, which forms the first part of the *Théorie Analytique* of Laplace.¹

¹ Lagrange, in vol. i. of the *Memoirs of the Society of Turin*, was the first who shewed that the investigation of the general term of a recurring series depends on the integration of a linear equation of finite differences. In vols. vi. and vii. of the *Memoires présentés à l'Académie des Sciences* of Paris, Laplace proposed a general method for the summation of recurring series by the integration of such equations, and in the latter volume gives a number of examples of their use in the more complicated questions in the theory of chances, amongst which is the problem enunciated in (108). The subject was afterwards resumed by Lagrange in the volume of the *Berlin Memoirs* for 1775, where he has given a more direct method than that of Laplace, for the integration of the class of equations in question, and also applied it to the solution of the principal problems proposed in the works of Montmort and Demoivre. A general solution of the problem in the text is given by Ampère in a Tract entitled *Considérations sur la Théorie Mathématique du Jeu*, (Lyons, 1802.)

SECTION IX.

OF THE MOST PROBABLE MEAN RESULTS OF NUMEROUS
DISCORDANT OBSERVATIONS, AND THE LIMITS OF PRO-
BABLE ERROR.

110. In the preceding section we have considered a class of questions which apply to events depending on constant causes, and supposed to be of such a nature that they necessarily happen or fail in each experiment, and have given formulæ by which approximate results can be obtained when the numbers involved are so large that they cannot be conveniently treated, or cannot be treated at all, by the ordinary methods of calculation. We come now to a more difficult problem, namely, to investigate the probable result of a large number of observations which have reference not to the simple occurrence or failure of a certain event, but to the magnitude of a thing, susceptible, within certain limits, of a very great or an infinite number of different values, equally or unequally probable, the chance of any particular value being also supposed to vary in each experiment. On account of its immediate application to the determination of the most probable values of astronomical and physical elements from the results of observation, this is, perhaps, in reference to practical utility, the most important question in the theory.

111. Let A represent a thing of any sort (as a line, or

an angle, or a function of any quantity) which may have every possible value within given limits, or which may be constant in itself, but of such a nature that its real magnitude can only be observed within certain limits of accuracy, and suppose a great number of observations to be made. The object is, in the first place, to assign the probability that the sum of the observed values shall fall within given limits, supposing the chances of the different values of A to be known *a priori*; and, in the second place, when the law of the chances is unknown, to determine from the observations themselves the most probable mean value of A , and also the limits within which there is a given amount of probability that the difference between such mean value, and the true but unknown value of A , shall be contained.

112. Let a and b be the limits of the possible values of A , x a value of A between a and b , and P the probability that the sum of the values of A given by h observations will be s exactly, s being a given quantity between ha and hb . Assume the values of A to be equidifferent, and multiples of a certain constant ϵ , and make

$$a\epsilon = a, \beta\epsilon = b, \sigma\epsilon = s, i\epsilon = x,$$

where a , β , and σ are whole numbers (which may be positive or negative), and i is also a whole number proportional to x , and varying between the limits $i = a$, and $i = \beta$, and which, therefore, may be positive or negative, or zero. If the different values of A are supposed to be equally probable, the chance of obtaining any given one of them, as x , in a single trial is unit divided by the number of possible values, or equal to $1 \div (\beta - a + 1)$; and if we assume an indeterminate quantity w , then (20) the number of combinations which give the sum of h values of A equal to $\sigma\epsilon$ is the coefficient of that term of the multinomial

$$(w^a + w^{a+1} + w^{a+2} \dots + w^\beta)^h,$$

(or of the developement of $(\Sigma w^i)^h$ from $i=a$ to $i=\beta$) in which the exponent of w is σ ; and consequently the probability P that the sum of the values of A will be s exactly is that coefficient divided by $(\beta-a+1)^h$.

113. If the chances of the different values of A are unequal, and also vary in each trial, let p_1 be the probability of the observed value of A being x in the first trial, p_2 the probability of its being x in the second, p_3 that of its being x in the third, and so on. Now when $h=1$, or when there is only a single trial, then $s=x=i\epsilon$, and we have $P=p^1$. If $h=2$, then, assuming $i\epsilon$ to be the value of A in the first trial, and $i'\epsilon$ its value in the second, (i and i' being any two numbers between a and β), the two observations may give the sum of the two values equal to $\sigma\epsilon$ in as many different ways as it is possible to satisfy the equation $i+i'=\sigma$; and consequently, according to the theory of combinations, P is the coefficient of that term of the product (arranged according to the powers of w) of the two series represented by $\Sigma p_1 w^{i\epsilon}$ and $\Sigma p_2 w^{i'\epsilon}$, in which the exponent of w is equal to $\sigma\epsilon$. In like manner, if $h=3$, then the sum of the observed values of A may be equal to $\sigma\epsilon$ in as many different ways as the equation $i+i'+i''=\sigma$ admits of different solutions, and consequently P is the coefficient of the term of the developement of the product $\Sigma p_1 w^{i\epsilon} \cdot \Sigma p_2 w^{i'\epsilon} \cdot \Sigma p_3 w^{i''\epsilon}$, in which the exponent of w is equal to $\sigma\epsilon$. Generally, when the number of observations is h , the probability P of the sum of the observed values of A being s , or $\sigma\epsilon$, exactly, is the coefficient of $w^{\sigma\epsilon}$ in the developement of the product

$$\Sigma p_1 w^{i\epsilon} \cdot \Sigma p_2 w^{i'\epsilon} \cdot \Sigma p_3 w^{i''\epsilon} \dots \Sigma p_h w^{i^{(h-1)}\epsilon},$$

the sums Σ including all values of i from $i=a$ to $i=\beta$:

Assume $w^e = e^{\theta\sqrt{-1}}$ (e being the base of the Napierian logarithms), and let the above product be denoted by X . We shall then have

$$X = \Sigma p_1 e^{i\theta\sqrt{-1}} \cdot \Sigma p_2 e^{i\theta\sqrt{-1}} \cdot \Sigma p_3 e^{i\theta\sqrt{-1}} \dots \Sigma p_h e^{i\theta\sqrt{-1}}.$$

Now since P is the coefficient of the term of the development of this product which contains the factor $e^{\sigma\theta\sqrt{-1}}$, if we conceive the developement effected we shall have

$$X = P e^{\sigma\theta\sqrt{-1}} + P' e^{\sigma'\theta\sqrt{-1}} + \&c.$$

a series in which all the terms are of the same form. Multiplying both sides of the equation by $e^{-\sigma\theta\sqrt{-1}}$, we get

$$X e^{-\sigma\theta\sqrt{-1}} = P + P' e^{(\sigma' - \sigma)\theta\sqrt{-1}} + \&c.$$

Now by a well known theorem in trigonometry, (ALGEBRA, art. 269), $e^{(\sigma' - \sigma)\theta\sqrt{-1}} = \cos (\sigma' - \sigma)\theta + \sqrt{-1} \sin (\sigma' - \sigma)\theta$; substituting therefore this value, and multiplying by $d\theta$, the equation becomes

$$X e^{-\sigma\theta\sqrt{-1}} d\theta = P d\theta + P' \{ \cos (\sigma' - \sigma)\theta + \sqrt{-1} \sin (\sigma' - \sigma)\theta \} d\theta + \&c.$$

The factor which multiplies P' in this equation will evidently become zero when integrated from $\ell = -\pi$ to $\ell = +\pi$, (π being the semicircumference to radius 1), the positive and negative elements of the integral being equal, and consequently destroying each other. The same thing also takes place with respect to the following terms, which are all of the same form. Integrating therefore between those limits, and observing that $\int d\theta = 2\pi$, we find

$$P = \frac{1}{2\pi} \int_{-\pi}^{+\pi} X e^{-\sigma\theta\sqrt{-1}} d\theta.$$

114. This value of P denotes the infinitely small chance that the sum of the values of A in h trials will be s exactly. Let μ and ν be two integer numbers between ha and $h\beta$, and let Q denote the probability that s will be comprised

between the two limits $\mu\epsilon$ and $\nu\epsilon$, (these limits being included between ha and hb); then Q will be found by substituting successively $\mu, \mu+1, \mu+2, \dots \nu$ for σ in the above value of P , and taking the sum of all the resulting terms. This substitution gives the following series multiplied by X under the sign of integration :

$$e^{-\mu\theta\sqrt{-1}} + e^{-(\mu+1)\theta\sqrt{-1}} + e^{-(\mu+2)\theta\sqrt{-1}} \dots + e^{-\nu\theta\sqrt{-1}}.$$

On multiplying the series now found by $e^{\frac{1}{2}\theta\sqrt{-1}} - e^{-\frac{1}{2}\theta\sqrt{-1}}$ ($= 2\sqrt{-1} \sin \frac{1}{2}\theta$), all the terms of the product, excepting the first and the last, destroy each other, and the sum of the terms becomes simply

$$e^{-(\mu-\frac{1}{2})\theta\sqrt{-1}} - e^{-(\nu+\frac{1}{2})\theta\sqrt{-1}};$$

therefore on making the substitution, and performing the multiplication now indicated, and dividing by $2\sqrt{-1} \sin \frac{1}{2}\theta$, we obtain for the value of Q the equation $Q=$

$$\frac{1}{4\pi\sqrt{-1}} \int_{-\pi}^{+\pi} X \left\{ e^{-(\mu-\frac{1}{2})\theta\sqrt{-1}} - e^{-(\nu+\frac{1}{2})\theta\sqrt{-1}} \right\} \frac{d\theta}{\sin \frac{1}{2}\theta}.$$

115. In order to simplify the expression for Q , let the number of possible values of A within the given limits be conceived to be infinite, in which case the constant ϵ becomes infinitely small, and therefore, since the limits are finite, μ and ν infinitely great. Let the following substitutions also be made :

$$\mu\epsilon = \psi - \delta, \quad \nu\epsilon = \psi + \delta, \quad \theta = \epsilon z,$$

δ being positive in order that ν may be greater than μ , agreeably to what has already been assumed. On substituting these expressions in the above equation, the limits of the new variable z will be \pm infinity; for ϵ having been supposed infinitely small, z must become infinitely great when $\theta = \pi$. Now since μ and ν are infinitely great, $\mu - \frac{1}{2}$ and $\nu + \frac{1}{2}$ become sensibly μ and ν , whence we have

$e^{-(\mu-\frac{1}{2})\theta\sqrt{-1}} - e^{-(\nu+\frac{1}{2})\theta\sqrt{-1}} = e^{-\psi z} (e^{\delta z\sqrt{-1}} - e^{-\delta z\sqrt{-1}})$
 $= e^{-\psi z} 2\sqrt{-1} \sin \delta z$. Again, by reason of $\theta = \epsilon z$, we have
 $d\theta = \epsilon dz$; and ϵ being infinitely small, θ must be a very
 small arc, therefore $\frac{1}{2}\theta$ may be taken for $\sin \frac{1}{2}\theta$, whence
 $d\theta \div \sin \frac{1}{2}\theta = 2dz \div z$. By means of these transformations
 the expression for Q becomes

$$Q = \frac{1}{\pi} \int_{-\infty}^{+\infty} X e^{-\psi z\sqrt{-1}} \sin \delta z \cdot \frac{dz}{z},$$

and denotes the probability that the sum of the h values of A will lie between $\psi \mp \delta$.

116. It is now necessary to assign a value to the product
 denoted by X . Since the number of possible values of A
 between a and b has been supposed infinite, the chance of
 obtaining any given one of them, as x , in a single trial, is
 infinitely small. Assuming this chance to be a function
 of x , and to vary in the different trials, let it be represented
 by $\phi_n x$ in respect of the n th trial. In order to preserve
 continuity in the values of A , this must be understood as
 signifying that $\phi_n x dx$ is the infinitely small chance that the
 value of A given by the n th observation will lie between
 x and $x + dx$. The function $\phi_n x$, therefore, represents the
law of the facility of the different values of A . It is posi-
 tive for all values of x between a and b , and vanishes for
 all values of x less than a or greater than b ; and it is im-
 portant to remark, that whatever number n may be, the
 integral $\int \phi_n x dx$ taken from $x=a$ to $x=b$ is always equal to
 unity; for since every observation gives a value of A be-
 tween a and b , the sum of all the probabilities in respect
 of each observation must be unity or certainty. From this
 assumption, then, we have $\phi_1 x dx = p_1$, $\phi_2 x dx = p_2$
 $\phi_n x dx = p_n$, whence the sums $\Sigma p_n e^{i\theta\sqrt{-1}}$ (113) are changed

into definite integrals; and therefore, since $\theta = \epsilon z$, $x = i\epsilon$, and consequently $i\theta = xz$, we obtain for the value of X ,

$$X = \int e^{zx} \sqrt{-1} \phi_1 x dx \cdot \int e^{zx} \sqrt{-1} \phi_2 x dx \dots \int e^{zx} \sqrt{-1} \phi_h x dx,$$

the limits of the integrals being $x=a$ and $x=b$.

By reason of $e^{zx} \sqrt{-1} = \cos zx + \sqrt{-1} \sin zx$, each of these integrals may be expressed in terms of the cosine and sine of zx . The n th, for instance, becomes $\int \phi_n x \cos zx dx + \sqrt{-1} \int \phi_n x \sin zx dx$. Now since $\int \phi_n dx = 1$, (from $x=a$ to $x=b$), and $\phi_n x$ can have only positive values, each of the integrals is less than 1; whence we may assume

$$\int \phi_n x \cos zx dx = R_n \cos r_n; \quad \int \phi_n x \sin zx dx = R_n \sin r_n;$$

R_n being a positive quantity, and r_n an angle having always a real value. This gives

$$\int e^{zx} \sqrt{-1} \phi_n x dx = R_n (\cos r_n - \sqrt{-1} \sin r_n) = R_n e^{r_n} \sqrt{-1};$$

whence substituting successively for n the numbers 1, 2, 3... h , and for the sake of brevity making

$$Y = R_1 \times R_2 \times R_3 \dots \times R_h,$$

$$y = r_1 + r_2 + r_3 \dots + r_h,$$

we get $X = Y e^{y} \sqrt{-1}$; and the expression for Q becomes

$$Q = \frac{1}{\pi} \int_{-\infty}^{+\infty} Y e^{(y-\psi z)} \sqrt{-1} \sin \delta z \cdot \frac{dz}{z}.$$

117. The integral in this last expression is equivalent to two others, namely

$$\int Y \cos(y-\psi z) \sin \delta z \cdot \frac{dz}{z} + \sqrt{-1} \int Y \sin(y-\psi z) \sin \delta z \cdot \frac{dz}{z}.$$

Now, on attending to the nature of the quantities represented by Y and y , it will be manifest that according as z is positive or negative, r_n and consequently y is positive or negative, while Y is positive in all cases, since R_n is always

positive. Hence $\cos(y - \psi z)$ is always positive, and the elements of the first of the above integrals having thus the same value and the same sign for the same value of z , whether z be positive or negative, the value of the integral from $-\infty$ to $+\infty$ is the double of its value from 0 to ∞ . On the other hand, since y and z have both the same sign, $\sin(y - \psi z)$ is positive or negative according as z is positive or negative, and the elements of the integral into which it enters being equal for $-z$ and $+z$, but having contrary signs, destroy each other, and the integral from $x = -\infty$ to $x = +\infty$ vanishes. The expression for Q is therefore transformed into

$$Q = \frac{2}{\pi} \int_0^{\infty} Y \cos(y - \psi z) \sin \delta z \cdot \frac{dz}{z}.$$

118. The formula now found cannot in general be integrated by any of the known methods, but in the present case the quantities denoted by Y and y are such that an approximate value of Q may be obtained, which will always be more nearly equal to the true value as h , the number of observations, is increased. On adding the squares of the two quantities represented by $R_n \cos r_n$ and $R_n \sin r_n$ we get

$$R_n^2 = (\int \phi_n x \cos zx \cdot dx)^2 + (\int \phi_n x \sin zx \cdot dx)^2.$$

If $z = 0$, this becomes $R_n = \int \phi_n x \cdot dx$, whence by (116), $R_n = 1$.

When z has a real value, then it may be shewn that R_n is less than 1; for let x' be any value of A different from x , then as x' can only vary from a to b , we have obviously, $\int \phi_n x' \cos zx' \cdot dx' = \int \phi_n x \cos zx \cdot dx$, and $\int \phi_n x' \sin zx' \cdot dx' = \int \phi_n x \sin zx \cdot dx$, and the above equation may be put under this form,

$R_n^2 = \int \phi_n x \cos zx \cdot dx \cdot \int \phi_n x' \cos zx' \cdot dx' + \int \phi_n x \sin zx \cdot dx \cdot \int \phi_n x' \sin zx' \cdot dx$,
whence

$$R_n^2 = \iint \phi_n x \phi_n x' \cos z(x - x') dx dx'.$$

Now, excepting the case in which $z=0$, this double integral is always less than $\iint \phi_n x \phi_n x' dx dx'$, or less than $(\int \phi_n x dx)^2$, and consequently R_n is less than $\int \phi_n x dx$, that is, less than unity. Since, then, it has been shewn that R_n is equal to unity when $z=0$, and less than unity for all other values of z , and since Y is a quantity of the order R_n^h , it follows that Y must diminish with great rapidity when z , or its equal $\theta \div \epsilon$ differs sensibly from 0, and even for very small values of z becomes insensible when h is a large number. We may therefore assume $Y=e^{-\theta^2}$, an expression which is equal to unity when $\theta=0$, and diminishes rapidly as θ is increased, and becomes zero when θ is infinite.

119. For the sake of abridging let us assume

$$k_n = \int x \phi_n x dx, \quad k'_n = \int x^2 \phi_n x dx, \quad k''_n = \int x^3 \phi_n x dx, \quad \&c.$$

(the integrals in respect of x being always from $x=a$ to $x=b$). From known formulæ we have

$$\cos zx = 1 - \frac{z^2 x^2}{2} + \frac{z^4 x^4}{2 \cdot 3 \cdot 4} - \&c.; \quad \sin zx = zx - \frac{z^3 x^3}{2 \cdot 3} + \&c.;$$

substituting these series for $\cos zx$ and $\sin zx$ in the integrals $\int \phi_n x \cos zx dx$ and $\int \phi_n x \sin zx dx$, and also $k_n, k'_n, k''_n, \&c.$, for the values they have now been assumed to represent, then, from (116) we have

$$R_n \cos r_n = 1 - \frac{z^2}{2} k'_n + \frac{z^4}{2 \cdot 4} k''_n - \&c.$$

$$R_n \sin r_n = z k_n - \frac{z^3}{2 \cdot 3} k''_n + \&c.$$

and it will be seen presently that all the terms involving higher powers of z than the cube may be neglected. Adding together the squares of these two equations, we get $R_n^2 = 1 - z^2 (k'_n - k_n^2) + z^4 f - \&c.$; whence

$$R_n = 1 - \frac{1}{2} z^2 (k'_n - k_n^2) + z^4 f' - \&c.;$$

f' being independent of z . On dividing the second by the

first, there results $\tan r_n = zk_n - \frac{1}{6}z^3k''_n + \frac{1}{2}z^5k_nk'_n - \&c.$;
whence by reason of $r_n = \tan r_n - \frac{1}{3} \tan^3 r_n + \&c.$,

$$r_n = zk_n - \frac{1}{6}z^3(k''_n - 3k_nk'_n + 2k_n^3) + \&c.$$

If, therefore, we make

$$c_n = \frac{1}{2}(k'_n - k_n^2), \quad g_n = \frac{1}{6}z^3(k''_n - 3k_nk'_n + 2k_n^3),$$

the values of R_n and r_n become respectively

$$R_n = 1 - z^2c_n + z^4f' - \&c. ; \quad r_n = zk_n - z^3g_n + \&c.$$

Now, by hypothesis (116) $Y = R_1 \times R_2 \times R_3 \dots R_h$;
therefore $\log Y = \Sigma \log R_n = \Sigma \log (1 - z^2c_n + z^4f' - \&c.) =$
 $-\Sigma \{z^2c_n - z^4(f' - \frac{1}{2}c_n) - \&c.\}$ (by reason of the formula
 $\log R_n = R_n - 1 - \frac{1}{2}(R_n - 1)^2 + \&c.$) But we have also assumed
(117) $Y = e^{-\theta^2}$; hence $\log Y = -\theta^2$, and consequently
 $\theta^2 = \Sigma \{z^2c_n - z^4(f' - \frac{1}{2}c_n) - \&c.\}$ In like manner, since
 $y = r_1 + r_2 + r_3 \dots + r_h = \Sigma r_n$, therefore $y = \Sigma zk_n - \Sigma z^3g_n +$
 $\&c.$ Now, the sums Σ include all values of c_n, k_n, g_n , from
 $n=1$ to $n=h$; let the mean values of those quantities,
therefore, be denoted by c, k, g , that is to say, let

$$\Sigma c_n = hc, \quad \Sigma k_n = hk, \quad \Sigma g_n = hg,$$

and make also $hf' = \Sigma(f' - \frac{1}{2}c_n)$, and we have $\theta^2 = z^2hc -$
 $z^4hf'' + \&c.$ By reverting the series the value of z is found

$$\text{in terms of } \theta ; \text{ namely } z = \frac{\theta}{\sqrt{hc}} + \frac{f''\theta^3}{2hc^2\sqrt{hc}} + \&c.$$

But the second term of this series, being divided by $h\sqrt{h}$,
and h being by supposition a large number, is very small in
comparison of the first, and may be neglected as insen-
sible. All the succeeding terms of the series are divided
by higher powers of h , and may therefore be rejected *a for-*
tiori. Confining the approximation, therefore, to terms of
the order $1 \div \sqrt{h}$, and rejecting all those into which h or
its powers enters as a divisor, we have $z = \theta \div \sqrt{hc}$, and
likewise $dz \div z = d\theta \div \theta$.

From (116) we have also $y = \Sigma r_n = \Sigma zk_n - z^3 \Sigma g_n + \&c.$,

therefore in consequence of the above transformations, $y = zhk - z^3hg + \&c.$; and on substituting for z its value just found in terms of θ , $y = kh\sqrt{(h \div c)} - g\theta^3 \div c\sqrt{(hc)}$, and consequently $y - \psi z = (hk - \psi)\theta \div \sqrt{(hc)} - g\theta^3 c \div \sqrt{(hc)}$. In order to deduce from this an expression for $\cos(y - \psi z)$, let u and v denote any two arcs, then by trigonometry, $\cos(u - v) = \cos u \cos v + \sin u \sin v$. Suppose v to be small, and let its cosine and sine be developed in series and substituted in this equation; it will become

$$\cos(u - v) = \cos u - \frac{v^2}{1.2} \cos u + \&c. + v \sin u - \frac{v^3}{2.3} \sin u + \&c.$$

whence, making $u = (hk - \psi)\theta \div \sqrt{(hc)}$, $v = g\theta^3 \div c\sqrt{(hc)}$, and rejecting as before terms of the order $1 \div h$, we have

$$\cos(y - \psi z) = \cos \left\{ (hk - \psi) \frac{\theta}{\sqrt{(hc)}} \right\} + \frac{g\theta^3}{c\sqrt{(hc)}} \sin \left\{ (hk - \psi) \frac{\theta}{\sqrt{(hc)}} \right\}.$$

If we now substitute the values of Y , z , dz , $\cos(y - \psi z)$ found in the last three paragraphs in the value of Q (117) we obtain the following expression in which the largest terms omitted are of the order $1 \div h$, and which therefore is more accurate as h is a higher number, viz.

$$Q = \frac{2}{\pi} \int_0^\infty e^{-\theta^2} \cos \left\{ (hk - \psi) \frac{\theta}{\sqrt{(hc)}} \right\} \sin \frac{\delta\theta}{\sqrt{(hc)}} \cdot \frac{d\theta}{\theta} \\ + \frac{2g}{\pi c\sqrt{(hc)}} \int_0^\infty e^{-\theta^2} \sin \left\{ (hk - \psi) \frac{\theta}{\sqrt{(hc)}} \right\} \sin \frac{\delta\theta}{\sqrt{(hc)}} \cdot \theta^2 d\theta.$$

120. As no restriction has yet been made with respect to the value of ψ , excepting that it is a mean between $\mu\epsilon$ and $\nu\epsilon$, and therefore included between ha and hb (115), let us now assume $\psi = hk$. This gives $\cos(hk - \psi) = 1$, and $\sin(hk - \psi) = 0$; and the equation becomes

$$Q = \frac{2}{\pi} \int_0^\infty e^{-\theta^2} \sin \frac{\delta\theta}{\sqrt{(hc)}} \cdot \frac{d\theta}{\theta},$$

which is the probability that the sum of the observed values of A will fall between $hk \pm \delta$.

121. The last step in this investigation is to reduce the integral now found to a known form, which may be accomplished as follows: Let u be a new variable, then by means of the trigonometrical formula $\cos u = \frac{1}{2}e^{u\sqrt{-1}} + \frac{1}{2}e^{-u\sqrt{-1}}$,

$$\int e^{-\theta^2} \cos u\theta. d\theta = \frac{1}{2} \int e^{-\theta^2+u\theta} \sqrt{-1} d\theta + \frac{1}{2} \int e^{-\theta^2-u\theta} \sqrt{-1} d\theta.$$

But $-\theta^2 + u\theta\sqrt{-1} = -\frac{1}{4}u^2 - (\theta - \frac{1}{2}u\sqrt{-1})^2$; assume therefore, $v = \theta - \frac{1}{2}u\sqrt{-1}$ (whence $dv = d\theta$), then

$$\frac{1}{2} \int e^{-\theta^2+u\theta} \sqrt{-1} d\theta = \frac{1}{2} \int e^{-\frac{1}{4}u^2 - v^2} dv = \frac{1}{2} e^{-\frac{1}{4}u^2} \int e^{-v^2} dv.$$

When $\theta=0$, then $v = -\frac{1}{2}u\sqrt{-1}$, and when θ is infinite, v is infinite; therefore, if the integral in respect of θ be taken from $\theta=0$ to $\theta=\infty$, the integral in respect of v must be taken from $v = -\frac{1}{2}u\sqrt{-1}$ to $v = \infty$.

If we now suppose u to be negative, we shall have in like manner $\frac{1}{2} \int e^{-\theta^2-u\theta} \sqrt{-1} d\theta = \frac{1}{2} e^{-\frac{1}{4}u^2} \int e^{-v^2} dv$, the limits in this case being from $v = +\frac{1}{2}u\sqrt{-1}$ to $v = \infty$. Hence

$$\int e^{-\theta^2} \cos u\theta. d\theta = \frac{1}{2} e^{-\frac{1}{4}u^2} (\int e^{-v^2} dv + \int e^{-v^2} dv).$$

But the sum of the two integrals on the right-hand side of this equation, the first being taken from $v = -\frac{1}{2}u\sqrt{-1}$ to infinity, and the second from $v = +\frac{1}{2}u\sqrt{-1}$ to infinity, is obviously the double of $\int e^{-v^2} dv$ from $v=0$ to $v = \infty$, or (96) equal to $\sqrt{\pi}$; and we have therefore

$$\int_0^\infty e^{-\theta^2} \cos u\theta. d\theta = \frac{1}{2} \sqrt{\pi} e^{-\frac{1}{4}u^2}.$$

Let both sides of the equation be multiplied by du , and integrated from $u=0$ to $u=\delta \div \sqrt{(hc)} = u'$; then observing that $\int \cos(u\theta) d\theta = \sin(u\theta) \div \theta$, we shall have

$$\int_0^\infty e^{-\theta^2} \sin \frac{\delta\theta}{\sqrt{(hc)}} \cdot \frac{d\theta}{\theta} = \frac{1}{2} \sqrt{\pi} \int_0^{u'} e^{-\frac{1}{4}u^2} du.$$

Comparing this equation with that in (110), we find $Q = (1 \div \sqrt{\pi}) \int_0^{u'} e^{-\frac{1}{2}u^2} du$. Now, let $u = 2t$, and let τ be what t becomes when $u = u' = \delta \div \sqrt{(hc)}$; then $\frac{1}{2}u^2 = t^2$, $du = 2dt$, $\delta \div \sqrt{(hc)} = 2\tau$, or $\delta = 2\tau\sqrt{(hc)}$, and we have, finally,

$$Q = \frac{2}{\sqrt{\pi}} \int_0^{\tau} e^{-t^2} dt, \text{ or, } Q = 1 - \frac{2}{\sqrt{\pi}} \int_{\tau}^{\infty} e^{-t^2} dt,$$

for the probability that s , the sum of the observed values of A , will be comprised between the limits $\psi - \delta$ and $\psi + \delta$, that is, between $hk \mp 2\tau\sqrt{(hc)}$; or, that the arithmetical mean of all the observations, namely $s \div h$, will lie between $k \mp 2\tau\sqrt{(c \div h)}$.

122. The expression now found for Q is that which in (96) was denoted by Θ , and of which the table gives the values corresponding to the different values of τ . The general result of the investigation is, therefore, that whatever be the nature of the function $\phi_n x$ which represents the law of the facility of the different values of A , if a large number of observations be made, the sum of the values of A , divided by the number of observations, approaches continually to a certain special quantity k (which is the true mean value of A) as the number of observations is increased, and that by multiplying the number of observations, a probability Θ may always be obtained, approaching as nearly to certainty as we please, that the difference between the arithmetical mean or average of the observations and the true mean value of A , will be comprised within limits which may be made as small as we please.

The analysis employed in the preceding articles, (113 to 121), for the purpose of establishing this very important result, belongs to Poisson, and is given in nearly the same form in the *Recherches sur la Probabilité des Juge-*

ments, chap. iv., and in the *Additions* to the *Connaissance des Temps* for 1832. We have preferred it to the method followed by Laplace in the *Théorie Analytique*, as being somewhat simpler and also more general.

123. In order that the limits $2\tau\sqrt{(hc)}$ may be real, it is necessary that the special quantity c be positive, a condition which has hitherto been assumed. Now, since $c = \Sigma c_n - h$, it is obvious that c will be positive if $c_n = \frac{1}{2}(k'_n - k_n^2)$ be positive. On writing for k'_n and k_n their values (119) we have

$$2c_n = \int x^2 \phi_n x dx - (\int x \phi_n x dx)^2,$$

the limits of the integrals being always from $x=a$ to $x=b$. But it is evident that no change will be made in the values of these definite integrals (the limits continuing the same), by substituting in them any other of the possible values of A , as x' . We have therefore $\int x' \phi_n x' dx' = \int x \phi_n x dx$, and since in all cases $\int \phi_n x' dx' = 1$, the above equation may be otherwise written

$$2c_n = \int x^2 \phi_n x dx \int \phi_n x' dx' - \int x \phi_n x dx \int x' \phi_n x' dx,$$

whence $2c_n = \iint \phi_n x \phi_n x' (x^2 - xx') dx dx'$,

or $2c_n = \iint \phi_n x \phi_n x' (x'^2 - x'x) dx dx'$.

Adding together the two last equations, there results

$$4c_n = \iint \phi_n x \phi_n x' (x - x')^2 dx dx',$$

a quantity which is necessarily positive, and can never be zero so long as x can have different values.

124. The special quantity k to which the average of the values of A continually approaches, is connected with the centre of gravity of the area of a curve by the following relation. Let x and y be the co-ordinates of a curve, of which the equation is $y = \phi_n x$; then the element of the area is $\phi_n x' dx$. But (116) $\phi_n x dx$ is the infinitely small probability that the value of A in the n th observation will lie between x and $x + dx$; therefore the element of the area of the curve

represents this probability, and the curve itself represents the law of the probability of the different values of A in respect of the n th trial. In like manner, the curve whose co-ordinates are x and $(1 \div h) \sum \phi_n x$, represents the law of the mean probability of A in respect of the whole series of observations. Now, if x_1 be the absciss of the centre of gravity of any curve whose co-ordinates are x and y , the well known formula of mechanics gives $x_1 = \int y x dx \div \int y dx$; therefore, applying this formula to the curve of the mean probability, and making the whole area ($\int y dx$ from $x=a$ to $x=b$) = 1, the absciss of the centre of gravity is $x_1 = (1 \div h) \sum \int x \phi_n x dx$. But this is the quantity denoted by k (119); hence the special quantity to which the average of a large number of observations indefinitely approaches is the absciss of the centre of gravity of the area of the curve which represents the law of the mean chances of A .

125. It has been assumed in the foregoing analysis that A is susceptible of an infinite number of values, increasing continuously from a to b . The results, however, are easily adapted to those cases in which the number of possible values of A is finite. Suppose A to be a thing susceptible of only λ different values, represented by $a_1, a_2, a_3, \dots, a_\lambda$, and let the chances of these values, which may be different in the different trials, be respectively $\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_\lambda$ in respect of the n th trial. Now, suppose $\phi_n x$ to be a discontinuous function, which vanishes for all values of x , of which the difference from one or other of the above values of A exceeds an infinitely small quantity ϵ ; then the whole integral $\int \phi_n x dx$ from $x=a$ to $x=b$, will be made up of a series of λ partial integrals $\int \phi_n x dx$ taken between the limits $a_i \pm \epsilon$, the sum of which will be unity, since one or other of the values of A must necessarily be given by the trial. But

the integral $\int \phi_n x dx$ between the limits $a_i \pm \epsilon$ is the expression of the chance that the value of A given in the n th trial will lie between $a_i \pm \epsilon$; whence for those limits $\int \phi_n x dx = \gamma_i$. Now the difference $x - a_i$ must be infinitely small, since it cannot exceed ϵ ; we may therefore substitute a_i for x , and a_i^2 for x^2 under the sign of integration, when the limits are $a_i \pm \epsilon$, so that for those limits we have $\int x \phi_n x dx = a_i \int \phi_n x dx = \gamma_i a_i$. On writing for i all the different numbers $1, 2, 3, \dots, \lambda$, and observing that the λ partial integrals thus formed make up the whole integral $\int x \phi_n x dx$ from $x=a$ to $x=b$, and that therefore their sum is k_n , we have, in respect of the n th trial,

$$k_n = \gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3 \dots + \gamma_\lambda a_\lambda.$$

In like manner, for $k'_n = \int x^2 \phi_n x dx$ (from a to b), we have

$$k'_n = \gamma_1 a_1^2 + \gamma_2 a_2^2 + \gamma_3 a_3^2 \dots + \gamma_\lambda a_\lambda^2;$$

so that the two special quantities k and k' become

$$k = (1 \div h) \Sigma (\gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3 \dots + \gamma_\lambda a_\lambda),$$

$$k' = (1 \div h) \Sigma (\gamma_1 a_1^2 + \gamma_2 a_2^2 + \gamma_3 a_3^2 \dots + \gamma_\lambda a_\lambda^2),$$

the sums Σ extending to all the h values of n , or to all the trials, the chances denoted by γ_1, γ_2 , &c. being supposed to vary in the different trials.

126. When the chances of the different values of A are equal and constant, then $\gamma_i = 1 \div \lambda$, and the above values of k and k' become

$$k = (1 \div \lambda) (a_1 + a_2 + a_3 \dots + a_\lambda),$$

$$k' = (1 \div \lambda) (a_1^2 + a_2^2 + a_3^2 \dots + a_\lambda^2),$$

so that k is the arithmetical mean of the possible values of A , and k' the mean of the squares of those values. On this hypothesis, therefore, k and k' may be computed *a priori*, and consequently the limits determined within which there is a given probability Θ that the average of h observations will fall, the limits being $k \pm 2\tau \sqrt{(c \div h)}$, where $c = \frac{1}{2}(k' - k^2)$.

When the chances of the different values of A are unequal, but constant in the different trials, then $k=k_n$, and $k'=k'_n$, and we have

$$\begin{aligned} k &= \gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3 \dots \dots + \gamma_\lambda a_\lambda, \\ k' &= \gamma_1 a_1^2 + \gamma_2 a_2^2 + \gamma_3 a_3^2 \dots \dots + \gamma_\lambda a_\lambda^2. \end{aligned}$$

In this case the special quantity k to which the average of the observed values continually approaches, is the sum of the possible values, each multiplied into its respective probability; and k' is the sum of the products of the squares of those values into their respective probabilities.

127. Resuming the consideration of the general formula in (121), we shall now give an example of its application when the function which represents the law of facility of the different values of A is supposed to be known *a priori*.

Of all the hypotheses which may be made respecting the law of facility, the simplest is that which supposes the chances of all the possible values of the thing observed to be equal, and to remain constant during the series of trials. This supposes $\varphi_n x = \varphi x = a$ constant; whence $\int \varphi x dx$, between the limits $x=a$ and $x=b$, becomes $(b-a) \varphi x$. But between those limits we have also $\int \varphi x dx = 1$; therefore $\varphi x = 1 \div (b-a)$. From this value of φx it is easy to deduce the special quantities k and k' . On the present hypothesis $k=k_n$ and $k'=k'_n$; therefore, the limits of the integral being $x=a$ and $x=b$, we have $k = \int x \varphi x dx = \int \frac{x dx}{b-a} = \frac{b^2 - a^2}{2(b-a)} = \frac{1}{2}(b+a)$, whence $k^2 = \frac{1}{4}(b+a)^2$. In like manner $k' = \int x^2 \varphi x dx$ becomes $\int \frac{x^2 dx}{b-a} = \frac{1}{3}(b^2 + ba + a^2)$; whence $c = \frac{1}{2}(k' - k^2) = \frac{1}{3}(b^2 + ba + a^2) - \frac{1}{3}(b+a)^2$. Hence by (121) we have

the probability Θ that the average value of A given by h observations, or the sum of the values of A divided by their number, will lie between

$$\frac{1}{2}(b+a) \mp 2\tau\sqrt{\left\{\frac{1}{6}(b^2 + ba + a^2) - \frac{1}{8}(b+a)^2\right\} \div \sqrt{h}}.$$

128. This formula may be applied to the following question. Of the comets which have been observed since the year 240 of our era, the parabolic elements of 138 have been computed, and the mean inclination of their orbits to the ecliptic is found to be $48^\circ 55'$. Now, supposing every possible inclination of an orbit to be equally probable, let the probability be demanded that the mean inclination of 138 orbits will not differ from 45° (the mean of the possible inclinations) more than 5° in excess or defect.

In this case the limits of the possible values of the phenomenon are 0 and 90° . We have therefore $a=0$, $b=90^\circ$, $h=138$, and the above limits of the error of the average, become $45^\circ \mp \tau \times 90^\circ \div \sqrt{6 \times 138}$. In order that the limits may not exceed 5° , we have to determine τ from the equation $\tau \times 90^\circ \div \sqrt{6 \times 138} = 5^\circ$, which gives $\tau = \frac{1}{3}\sqrt{23}$; whence $\tau=1.6$ very nearly. The tabular value of Θ corresponding to $\tau=1.6$ is $.97635$, or nearly $\frac{4}{5}$; and the odds are therefore 41 to 1 that on the supposition of all inclinations being equally probable, the mean inclination of 138 comets would fall between $45^\circ \mp 5^\circ$, that is, between 40° and 50° . The mean of the inclinations actually computed falls within those limits (being $48^\circ 45'$); there is therefore a very great probability that whatever may be the nature of the unknown causes which determine the positions of the cometary orbits, it is not such as to render different inclinations unequally probable.

If the question had been to assign the limits within which it is as probable that the mean of the inclinations will fall

as not, we should have had $\Theta = \frac{1}{2}$, and consequently (from the table) $\tau = .476936$, and the limits would have been $45^\circ \mp 90^\circ \times .476936 \div \sqrt{(6 \times 138)}$, which is found on calculation to be $45^\circ \mp 1^\circ.5$. On the supposition, therefore, that all inclinations are equally probable, it is one to one that the mean of 138 inclinations will fall between $43\frac{1}{2}^\circ$ and $46\frac{1}{2}^\circ$, or at least not exceed those limits.

129. On the same hypothesis of an equal probability of all possible values, if we suppose the mean value of A to be 0, we have then $a = -b$, and ϕx becomes $1 \div 2a$, whence the limits corresponding to a given value of Θ (127) become $0 \mp 2\tau b : \sqrt{(6h)}$. Let $\Theta = \frac{1}{2}$, whence $\tau = .476936$, and suppose $h = 600$. With these values the limits become $0 \mp .016b$ nearly; that is to say, it is an even wager that the average of 600 observations will not differ from the true mean value of A more than the sixteen-thousandth part of a or b , what is the greatest possible difference.

130. As a second hypothesis, suppose the chance of a given value of A to decrease uniformly as the magnitude increases from 0 to $\pm a$; then ϕx will be found as follows: Let $\phi x = \beta$ when $a = 0$; we have then by the hypothesis $\phi x : \beta = (a - x) : a$, whence $\phi x = (a - x)\beta \div a$, and consequently $\int \phi x dx = \beta x - \beta x^2 \div 2a$, which, from $x = 0$ to $x = +a$, becomes $\frac{1}{2}\beta a$. But $\int \phi x dx$ from $x = -a$ to $x = +a$ is 1 (errors beyond those limits being supposed impossible), therefore from $x = 0$ to $x = +a$, $\int \phi x dx = \frac{1}{2}$, and consequently $\frac{1}{2}\beta a = \frac{1}{2}$, or $\beta = 1 \div a$. Hence $\phi x = (a - x) \div a^2$, from which the value of c is easily deduced, that of h being 0, as in the former case.

131. Although the function ϕx which represents the law of facility of the different values of A is in general unknown, its form may be assigned if we assume that it is subject

to certain conditions, which, from the nature of the thing, must be very nearly, if not absolutely true, in most practical cases: 1st, That the chance of an error diminishes as the magnitude of the error increases, and for errors beyond a certain limit vanishes altogether; and, 2d, that positive and negative errors, of equal magnitude, are equally probable. The last condition is equivalent to the assumption that the average of the observed values is the true mean value. For simplification, we suppose the chance of an error of a given magnitude to remain constant in all the trials.

132. Let $x, x', x'', \&c.$ be a series of values of A , the sum of which is s , and the number h , and make $m = s \div h$, then m is the arithmetical mean or average, which by hypothesis is the true value of the phenomenon A . Let $x - m = \Delta$, $x' - m = \Delta'$, $x'' - m = \Delta''$, $\&c.$, so that $\Delta, \Delta', \Delta'', \&c.$ are the errors of $x, x', x'', \&c.$ Now, the most probable single error is 0; and the probability of obtaining an error of a given magnitude Δ in any observation is obviously the same as that of obtaining a given value of x ; therefore $\phi x = \phi(x - m) = \phi \Delta$; so that $\phi \Delta$ is the probability of a single error being exactly Δ . In like manner, the probability of an error $= \Delta'$ is $\phi \Delta'$; and if we take P to denote the probability of a given system of errors, $\Delta, \Delta', \Delta'', \&c.$, then the errors being supposed independent of each other, we have (7)

$$P = \phi \Delta \cdot \phi \Delta' \cdot \phi \Delta'', \&c.$$

Let this system be assumed to be the most probable result of the observations, then P is a maximum, and its differential co-efficient zero. Taking the logarithms of both sides of the equation, differentiating, and making $d \log. \phi \Delta = \phi' \Delta d\Delta$, and $dP \div d\Delta = 0$, we obtain

$$0 = \phi' \Delta + \phi' \Delta' + \phi' \Delta'' +, \&c.,$$

an equation which may be otherwise written

$$0 = \Delta \frac{\phi' \Delta}{\Delta} + \Delta' \frac{\phi' \Delta'}{\Delta'} + \Delta'' \frac{\phi' \Delta''}{\Delta''} +, \&c.$$

This is the conditional equation of the most probable system of errors. But the hypothesis of the average being the true value, furnishes this other equation,

$$0 = (x - m) + (x' - m) + (x'' - m) +, \&c.$$

or, which is the same, $0 = \Delta + \Delta' + \Delta'' +, \&c.$; and on comparing this with the above conditional equation, it is evident that they can only be both true simultaneously on the supposition of $\frac{\phi' \Delta}{\Delta} = \frac{\phi' \Delta'}{\Delta'} = \frac{\phi' \Delta''}{\Delta''} =, \&c.$ Hence it follows

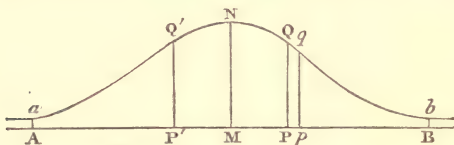
that $\phi' \Delta \div \Delta$ is independent of any particular value of Δ , or is equal to a constant, which we shall call K . We have then

$$\frac{\phi' \Delta}{\Delta} = \frac{d \cdot \log. \phi \Delta}{\Delta d \Delta} = K.$$

The integral of this expression is $\log. \phi \Delta = \frac{1}{2} K \Delta^2 + \text{const.}$, which, making the last constant $= \log. H$, and passing to numbers, gives $\phi \Delta = H e^{\frac{1}{2} K \Delta^2}$. It now only remains to determine the two constants H and K . With respect to K , as we suppose the most probable value of Δ to be 0, and that $\phi \Delta$ diminishes as Δ increases, it is obvious that K must be negative. Assume $\frac{1}{2} K = -\gamma$, and the formula becomes $\phi \Delta = H e^{-\gamma \Delta^2}$. For the determination of H we have the equation $\int \phi \Delta d \Delta = 1$, the limits of the integral being $-a'$ and $+a'$, where $a' = \frac{1}{2}(b-a)$, a and b being the limiting values of x . But it is to be observed, that as all values of Δ exceeding the limits $\pm a'$ are supposed to be impossible, or at least to be so improbable that it is unnecessary to take account of them, the value of the integral $\int \phi \Delta d \Delta$ from $\Delta = -a'$ to $\Delta = +a'$ will not be altered by extending the limits from $-$ infinity to $+$ infinity. We have therefore

$\int_{-\infty}^{+\infty} \phi \Delta d\Delta = 1$. Let $\Delta = t \div \sqrt{\gamma}$, then $d\Delta = dt \div \sqrt{\gamma}$, and $\phi \Delta = H e^{-\gamma \Delta^2} = H e^{-t^2}$, on substituting which in the last equation, and observing that from $t = -\infty$ to $t = +\infty$ we have $\int e^{-t^2} dt = \sqrt{\pi}$ (96), we find $(H \div \sqrt{\gamma}) \sqrt{\pi} = 1$, and $H = \sqrt{(\gamma \div \pi)}$. Whence, finally, $\phi \Delta = \sqrt{(\gamma \div \pi)} e^{-\gamma \Delta^2}$.

133. The general properties of the function now found may be illustrated by means of a curve line. Let aNb be



a curve of which $\phi \Delta$ is the ordinate corresponding to the absciss Δ . Let AB be its axis, and MN its greatest ordinate. Suppose the origin to be placed at M , draw PQ an ordinate at any point P , and pq indefinitely near to PQ , and make $MB = a'$, $MA = -a'$, $MP = \Delta$, and $PQ = \phi \Delta$; then as was shewn in (124), $PQqp$, the element of the area, represents the chance of an error lying between Δ and $\Delta + d\Delta$, that is, of an error greater than MP but less than Mp . Now, if $\phi \Delta = \sqrt{(\gamma \div \pi)} e^{-\gamma \Delta^2}$ the function will not be changed by changing Δ into $-\Delta$; therefore $\phi \Delta = \phi(-\Delta)$, and the curve is symmetrical on both sides of MN , as it obviously ought to be according to the hypothesis; for on making $MP' = MP$, then positive and negative errors of equal magnitude being equally probable, we must have $P'Q' = PQ$. Again, since $e^{-\gamma \Delta^2}$ diminishes rapidly as Δ increases, the curve at a short distance from MN must approach very near to its axis AB ; but as the function only vanishes when Δ is infinite, the curve will not meet the axis at any finite distance from MN . This curve, therefore, can only represent approximately the law of

facility, inasmuch as it is supposed that errors beyond a certain limit are impossible; but on account of the rapid diminution of the ordinate at a short distance from MN, the chance of an error exceeding a small value of Δ , as MB, becomes insensible. Hence the limits of the integrals in respect of Δ may be extended without sensibly altering their values from $\Delta = \pm a'$ to $\Delta = \pm \infty$.

134. It is now necessary to find the special quantities k , k' , and c . Substituting Δ for x , and observing that as the law of the chances is here supposed to remain constant, we have $k = k_n$, $k' = k'_n$, the formulæ in (119) become $k = \int \Delta \phi \Delta d\Delta$, $k' = \int \Delta^2 \phi \Delta d\Delta$. Hence on making $\phi \Delta = \sqrt{(\gamma \div \pi)} e^{-\gamma \Delta^2}$, we have

$$k = \sqrt{\frac{\gamma}{\pi}} \int \Delta e^{-\gamma \Delta^2} d\Delta = \sqrt{\frac{\gamma}{\pi}} \cdot \frac{1}{2\gamma} \int d. e^{-\gamma \Delta^2} = \frac{1}{2\sqrt{(\gamma \pi)}} e^{-\gamma \Delta^2}.$$

When Δ becomes infinite, this becomes 0, therefore from $\Delta = -\infty$ to $\Delta = +\infty$, $k = 0$. This is an obvious consequence of the symmetry of the curve, for the centre of gravity is necessarily in the straight line MN.

With respect to k' we may proceed thus. We have $k' = \int \Delta^2 \phi \Delta d\Delta = \sqrt{(\gamma \div \pi)} \int \Delta^2 e^{-\gamma \Delta^2} d\Delta$. But from the principles of the differential calculus,

$$d. \Delta e^{-\gamma \Delta^2} = e^{-\gamma \Delta^2} d\Delta - 2\gamma \Delta^2 e^{-\gamma \Delta^2} d\Delta,$$

therefore, integrating and transposing,

$$\int \Delta^2 e^{-\gamma \Delta^2} d\Delta = -\frac{1}{2\gamma} \Delta e^{-\gamma \Delta^2} + \frac{1}{2\gamma} \int e^{-\gamma \Delta^2} d\Delta.$$

Now, from $\Delta = -\infty$ to $\Delta = +\infty$, the term of this equation which is not under the sign of integration vanishes, and $\int e^{-\gamma \Delta^2} d\Delta = \sqrt{(\pi \div \gamma)}$ (from (96), on substituting t^2 for $\gamma \Delta^2$), therefore $\int \Delta^2 e^{-\gamma \Delta^2} d\Delta = (1 \div 2\gamma) \sqrt{(\pi \div \gamma)}$; and consequently $k' = 1 \div 2\gamma$.

In (119) we assumed $c = \frac{1}{2}(h' - k^2)$; therefore in the present case $c = \frac{1}{2}h'$, whence $c = 1 \div 4\gamma$, or $\gamma = 1 \div 4c$.

135. The expressions which have now been found for the function which represents the probability of an error, and the limits corresponding to an assigned degree of probability, are given in terms of the indeterminate constant γ (or c), which depends on the nature of the observation, and therefore, where instruments are requisite, on the goodness of the instrument and the skill of the observer. This constant is called by Laplace the *modulus* of the law of facility. It cannot, in general, be assigned *a priori*; but if we assume that positive and negative departures from the mean are alike probable, which is the most plausible hypothesis the nature of the thing admits of, an approximation to its value, in respect of observations of a given kind, may be deduced with great probability from the results of a large series of observations of the same kind already made. We now proceed to give the analysis by which this is accomplished, following the method of Poisson. The approximation is carried only to quantities of the order $1 \div \sqrt{h}$; terms having h for a divisor are neglected on account of their smallness, h being supposed a large number.

136. In the expression for Q in (119), suppose $\psi = \delta$, and consequently $\psi - \delta = 0$, $\psi + \delta = 2\delta$, and write also z for $\theta \div \sqrt{hc}$; the equation then becomes

$$Q = \frac{2}{\pi} \int_0^\infty e^{-\theta^2} \cos(hkz - \delta z) \sin \delta z \cdot \frac{\delta \theta}{\theta} \\ + \frac{2g}{\pi c \sqrt{hc}} \int_0^\infty -\theta^2 \sin(hkz - \delta z) \sin \delta z \cdot \theta^2 d\theta,$$

and Q is the probability that s , the sum of the values of A given by all the observations, will lie between 0 and 2δ . If therefore, we suppose δ to be variable, the differential of

this expression taken with respect to δ , will express the infinitely small chance of the sum of the values being 2δ exactly. Differentiating, and observing that if u and v denote any two arcs, the trigonometrical formulæ give

$$\begin{aligned}\sin(u-v) \sin v + \cos(u-v) \cos v &= \cos(2v-u), \\ -\cos(u-v) \sin v + \sin(u-v) \cos v &= -\sin(2v-u),\end{aligned}$$

we shall find

$$\begin{aligned}\frac{dQ}{d\delta} d\delta &= \frac{2d\delta}{\pi} \int_0^\infty e^{-z^2} \cos(2\delta z - h k z) \frac{z d\theta}{\theta} \\ &\quad - \frac{2g d\delta}{\pi c \sqrt{hc}} \int_0^\infty e^{-\theta^2} \sin(2\delta z - h k z) z \theta^2 d\theta.\end{aligned}$$

Let t be a variable quantity, and assume $2\delta = h k + 2t \sqrt{hc}$, whence $d\delta = dt \sqrt{hc}$, and let the corresponding value of $\frac{dQ}{d\delta} d\delta$ be denoted by $q dt$, we shall have, on substituting

$$\text{these values, and replacing } z \text{ by } \theta \div \sqrt{hc}, \quad q dt = \frac{2dt}{\pi} \int_0^\infty e^{-\theta^2} \cos(2t\theta) d\theta - \frac{2g dt}{\pi c \sqrt{hc}} \int_0^\infty e^{-\theta^2} \sin(2t\theta) \theta^3 d\theta.$$

The two integrals in this equation are found from the formula in (121). Writing $2t$ for u , that formula gives

$$\int_0^\infty e^{-\theta^2} \cos(2t\theta) d\theta = \frac{1}{2} \sqrt{\pi} \cdot e^{-t^2};$$

and if this last equation be differentiated in respect of t , three times in succession, the result will be

$$\int_0^\infty e^{-\theta^2} \sin(2t\theta) \theta^3 d\theta = \frac{1}{4} \sqrt{\pi} (3te^{-t^2} - 2t^3 e^{-t^2});$$

whence, if we make $V = \frac{g}{2c \sqrt{hc}} (3t - 2t^3)$, we shall have

$$q dt = (1 \div \sqrt{\pi}) (1 - V) e^{-t^2} dt,$$

where V is a quantity containing only uneven powers of t , and of the order $1 \div \sqrt{h}$, so that when multiplied by another of the same order, the product will be of the order $1 \div h$,

and will therefore be rejected in the present approximation. This value of qdt is the probability that s will be precisely 2δ or $hk + 2t\sqrt{hc}$, or it is the infinitely small probability of the equation

$$s = hk + 2t\sqrt{hc}.$$

137. In order to apply this result to the determination of the probable limits in terms of observations actually made, it is necessary to remark that the analysis by means of which it has been obtained is grounded on the very general supposition that the thing to be measured may be any function whatever of the quantity observed; for the infinitely small chance of a particular value of the function is evidently the same as that of the corresponding value of the quantity, and is consequently $\phi_n x dx$. Let X therefore be a function of x , and let K, C, T , be what h, c, t become when X is substituted for x , the above equation then becomes

$$\Sigma X = hK + 2T\sqrt{hC},$$

the symbol Σ including all the h values of X ; and the probability of this equation is an expression of the same form as that which is represented by qdt .

138. Hitherto no restriction has been made with respect to ϕx ; we now introduce the hypothesis that positive and negative departures from the mean of equal magnitude are equally probable, and consequently that the curve representing the law of facility is symmetrical, but shall suppose the chances of a particular value, or a particular error, to vary in the different trials. Let the origin be transferred to the centre of gravity, the absciss of which $= k$, and let $x - k = \Delta$, $x' - k = \Delta'$ &c. We have then by (132) $\phi x = \phi \Delta$, $\int x \phi x dx = \int \Delta \phi \Delta d\Delta$, $\int x^2 \phi x dx = \int \Delta^2 \phi \Delta d\Delta$, the integration in respect of Δ being from $-\infty$ to $+\infty$. The special

quantities k and k' then become $k=(1\div h)\Sigma f\Delta\phi_n\Delta d\Delta$
 $\equiv 0$, $k'=(1\div h)\Sigma f\Delta^2\phi_n\Delta d\Delta$, whence

$$c=(1\div 2h)\Sigma f\Delta^2\phi_n\Delta d\Delta.$$

The object is now to eliminate $\phi_n\Delta$, and determine c in terms of the observations.

139. Let $\lambda_n=x$ be the observed value of A in the n th observation, then $\lambda_n-k=\Delta$ is the true error of the observation. Let the function denoted by X in (137) be $(\lambda_n-k)^2=\Delta^2$, and the corresponding value of K (since in this case, $K=(1\div h)\Sigma fX\phi_nxdx$) becomes $K=(1\div h)\Sigma f\Delta^2\phi_n\Delta d\Delta$. Comparing this with the value of c found above, we have $K=2c$; therefore on substituting these values of X and K in the equation (137), and assuming t' and c' to be the values of T and C when $X=(\lambda_n-k)^2$, we get $\Sigma(\lambda_n-k)^2=2hc+2t'\sqrt{(hc')}$, whence

$$c=(1\div 2h)\Sigma(\lambda_n-k)^2-t'U' \quad (1)$$

(U being a quantity of the order $1\div\sqrt{h}$); and the probability of this equation is

$$q'dt'=(1\div\sqrt{\pi})(1-V')e^{-t'^2}dt',$$

where V' is a function containing only uneven powers of t' , and of the order $1\div\sqrt{h}$.

In the equation (137) suppose $X=x=\lambda_n$, and let t'' and c'' be the corresponding values of T and C , then since on this supposition $K=k$, the equation becomes $\Sigma\lambda_n=hk+2t''\sqrt{(hc'')}$, whence

$$k=(1\div h)\Sigma\lambda_n-t''U'', \quad (2)$$

(U'' being of the order $1\div\sqrt{h}$); and the probability of this equation is

$$q''dt''=(1\div\sqrt{\pi})(1-V'')e^{-t''^2}dt'',$$

where V'' , like V' and V , contains only uneven powers of t'' and is of the order $1\div\sqrt{h}$.

140. The two equations (1) and (2) may be regarded

as two distinct events, having the respective probabilities now assigned to them, and therefore the probability of their being true simultaneously is the product of their respective probabilities, and is accordingly (neglecting the product $V'V''$ which is a quantity divided by h),

$$q'q''dt'dt''=(1\div\pi)(1-V'-V'')e^{-t^2}e^{-t''^2}dt'dt''.$$

Let the value of h given by equation (2) be substituted in (1), and the expression now given will accordingly be the probability of the resulting equation, namely,

$$c=\frac{1}{2h}\Sigma(\lambda_n-\frac{1}{h}\Sigma\lambda_n+t'U')^2-t''U''.$$

Let $m=(1\div h)\Sigma\lambda_n$, then m is the average or arithmetical mean of the observed values, and λ_n-m the *reputed* error of the observation. The last equation will then become $c=(1\div 2h)\Sigma(\lambda_n-m+t'U')^2-t''U''$; or, rejecting $(t'U')^2$ which is of the order $1\div h$,

$$c=(1\div 2h)\Sigma\{(\lambda_n-m)^2+2(\lambda_n-m)t'U'\}-t''U''.$$

For the sake of abridging let us also assume

$$\mu=(1\div h)\Sigma(\lambda_n-m)^2, \nu=(1\div h)\Sigma(\lambda_n-m)t',$$

so that μ is the mean of the squares of the errors, or *mean square* of the errors, and the equation becomes

$$c=\frac{1}{2}\mu+\nu U'-t''U'', \quad (3)$$

the probability of which is $q'q''dt'dt''$.

141. Now, by (121), we have the probability \ominus that $s\div h$, or $\Sigma\lambda_n\div h=m$ the arithmetical mean of the observed values of A , will fall within the limits $h\mp 2\tau\sqrt{(c\div h)}$. Substituting in those limits the above value of c , and observing that $(\frac{1}{2}\mu+\nu U'-t''U'')^{\frac{1}{2}}=\sqrt{(\frac{1}{2}\mu)+N(\nu U'-t''U'')+&c.}$, and that U' and U'' being of the order $1\div\sqrt{h}$, when divided again by \sqrt{h} are to be rejected, the limits become

$$h\mp 2\tau\sqrt{(\frac{1}{2}\mu\div h)}, \text{ or } h\mp \tau\sqrt{(2\mu\div h)},$$

and the probability of these being the true limits is \ominus mul-

multiplied into the probability of the equation $c = \frac{1}{2}\mu + \nu U' - t''U''$; and is therefore (140)

$$(1 \div \pi) \Theta(1 - V' - V'') e^{-t'^2} e^{-t''^2} dt' dt''.$$

142. The expression now obtained is the infinitely small probability of the limits $k \pm \tau \sqrt{(2\mu \div h)}$ of the average m , in respect of the particular value of s , for which we have deduced the equation (3). But for every value of s between the limits 0 and 2δ , there will be an equation corresponding to (3); therefore, in order to have the whole probability of those limits, the integral of the expression must be found for all values of t' and t'' . From the nature of the expressions $e^{-t'^2}$ and $e^{-t''^2}$ as well as the consideration that errors beyond a certain magnitude, though possible, are wholly improbable, it is evident that the integration may be extended without sensible error from $-\infty$ to $+\infty$; and since the functions V' and V'' contain only uneven powers of t' and t'' , the terms into which they enter, disappear in the integrations between those limits. (See Lacroix, *Calcul Diff. et Integral*, tom. iii. p. 506). Now, from $t' = -\infty$ to $t' = +\infty$ we have (96) $\int e^{-t'^2} dt' = \sqrt{\pi}$; and $\int e^{-t''^2} dt'' = \sqrt{\pi}$; therefore

$$\frac{1}{\pi} \Theta \iint (1 - V' - V'') e^{-t'^2} e^{-t''^2} dt' dt'' = \Theta.$$

The result of the preceding analysis is therefore that on the hypothesis of positive and negative errors of equal magnitude being equally probable, and on rejecting terms divided by h (the number of the observations may be always so great as to render such terms insensible), we may substitute $\frac{1}{2}\mu$ for c in the limits of the error to be apprehended, without sensibly altering the probability, and consequently

there is the probability $\Theta = \frac{2}{\sqrt{\pi}} \int_0^\tau e^{-t^2} dt$ that the true

mean value h of the phenomenon A will lie between the limits

$$m \pm 2\tau \sqrt{(\frac{1}{2}\mu \div h)}, \text{ or } m \pm \tau \sqrt{(2\mu \div h)},$$

which contain only quantities given by observation.

On this hypothesis we have also (138) $c = (1 \div 2h) \sum f \Delta^2 \phi_n \Delta d \Delta$, or, supposing the law of facility to remain constant during the trials, $c = \frac{1}{2} f \Delta^2 \phi \Delta d \Delta$, therefore $\mu = f \Delta^2 \phi \Delta d \Delta$; that is to say, the mean of the squares of the actual errors may be taken for the sum of the products of the squares of the possible errors multiplied by their respective probabilities. It is important to remark that as the observations become more numerous, the quantity μ , the mean of the squares of the errors, converges more and more to a constant quantity, and finally becomes independent of the number of observations.

143. The limits now found may be otherwise expressed. By hypothesis, $m = (1 \div h) \sum \lambda_n$ = the arithmetical mean of the observed values, and $\mu = (1 \div h) \sum (\lambda_n - m)^2$ = the mean of the squares of the reputed errors. Now $(\lambda_n - m)^2 = \lambda_n^2 - 2\lambda_n m + m^2$, and $(1 \div h) \sum 2\lambda_n m = 2m (1 \div h) \sum \lambda_n = 2m^2$; therefore $\mu = (1 \div h) \sum \lambda_n^2 - m^2$, that is to say, the mean of the squares of the observations minus the square of the mean. Hence the limits, corresponding to a given probability Θ , of the difference between the average of all the observations and the true value, are expressed by either of these formulæ

$$\pm \tau \sqrt{\{2 \times \text{mean square of errors} \div h\}},$$

$\pm \tau \sqrt{\{2 \times \text{mean square of obs.} - (\text{mean of obs.})^2\} \div \sqrt{h}}$,
 h being the number of observations, and the relation between Θ and τ being given by the table. Generally speaking, the first of these formulæ is the most convenient for calculation.

144. Let l be the limit of the error to be feared in taking the average of the observations as the true result, then $l = \tau \sqrt{(2\mu \div h)}$, and $\tau = l \sqrt{(h \div 2\mu)}$. Now when τ is constant, that is, for a given probability Θ , the determination will be more exact in proportion as l is a smaller number, and the precision will therefore be proportional to $\sqrt{(h \div 2\mu)}$. Hence $\sqrt{(h \div 2\mu)}$ is called by Gauss the *measure of the precision* of the determination. Suppose two series of observations to have been made for the determination of an element, the comparative accuracy of the results will depend on two things, the number of observations in each series, and the amount of the squares of the errors in each. If the number of observations is the same in both series, the precision of each result will be inversely as the square root of the sum of the squares of the errors, and the presumption of accuracy is in favour of that result with respect to which the sum of the squares of the errors is less than in the other. On the other hand, if the mean square of the errors is the same in both series, then the observations are alike good in both, and their relative values of the two results are directly as the square roots of the number of observations in each series. Hence, in order that one determination may be twice as good as another, it must be founded on four times the number of equally good observations. These considerations are very important, in comparing tables of mean values of whatever kind, for example, of the probabilities of life at the different ages, and in estimating risks which depend upon them.

145. Astronomers employ the terms, *weight*, *probable error*, and *mean error*, of a result, to denote certain functions of μ , the mean square of the errors. The square of the quantity which measures the precision of the result, is

called the *weight* of the determination. Denoting the weight by w , we have therefore

$$w = h \div 2\mu = h^2 \div 2\Sigma(\lambda_n - m)^2,$$

or the weight is equal to the square of the number of observations divided by twice the sum of the squares of the errors. Substituting this in the expression of the limits, we have $l = \tau \div \sqrt{w}$, and $\tau = l\sqrt{w}$; that is to say, for a given probability Θ , the limits of the error to be apprehended in taking the average as the true result are reciprocally proportional to the square root of the weight. When observations of different kinds, or results deduced from observation, are compared with each other, their relative weights (supposing the number of observations the same) are inversely as μ , and are expressed numerically by taking the weight of a certain series of observations as the unit of weight.

146. The *probable error* of the determination is that which corresponds to the probability $\Theta = \frac{1}{2}$. For $\Theta = \frac{1}{2}$ we have $\tau = .476936$; whence $\tau\sqrt{2} = .674489$, and the formula $m \pm \tau\sqrt{(2\mu \div h)}$ becomes $m \pm .674489\sqrt{(\mu \div h)}$; whence

$$\text{probable error} = .674489\sqrt{(\mu \div h)}.$$

147. The *mean error* of the result of a large number of observations may be deduced from the general formula in (136) as follows. That formula gives $qdt = (1 \div \sqrt{\pi})(1 - V)e^{-t^2}dt$ for the probability that the sum of the observed values will be $2\delta = hk + 2t\sqrt{(hc)}$ exactly. Dividing the sum by h , qdt is also the probability that the average value given by all the observations will be exactly $k + 2t\sqrt{(c \div h)}$. Now, on the hypothesis that positive and negative departures from the mean are equally probable, and supposing the origin of the co-ordinates to be transferred to the centre of gravity of the curve of mean proba-

bility, we have $h=0$, and $qdt=(1\div\sqrt{\pi})(1-V)e^{-t^2}dt$ is the infinitely small chance of the average error being $2t\sqrt{(c\div h)}$ exactly. Multiplying therefore this error into the chance of its taking place, and integrating the product from $t=0$ to $t=\infty$, we shall have the *mean* error, or *mean risk* of all the possible average errors affected with the positive sign. Now, observing that V represents a quantity divided by \sqrt{h} , and therefore when multiplied by $2t\sqrt{(c\div h)}$ becomes of the order $1\div h$, and may consequently be rejected, the product of the average error $2t\sqrt{(c\div h)}$ into its probability is $2\sqrt{(c\div \pi h)} \times te^{-t^2}dt$; and since $\int te^{-t^2}dt = \frac{1}{2}\int d.e^{-t^2} = \frac{1}{2}e^{-t^2}$, which from $t=0$ to $t=\infty$ becomes simply $\frac{1}{2}$, the integral of the above product from $t=0$ to $t=\infty$ is $\sqrt{(c\div \pi h)}$. Substituting for c its value (142) $=\frac{1}{2}\mu$, this result becomes $\sqrt{(\mu\div 2\pi h)}$; whence on computing $\sqrt{(1\div 2\pi)}$ we obtain

$$\text{mean error of series} = .398942\sqrt{(\mu\div h)}.$$

This is the mean error or mean risk in respect of positive errors alone, or on the supposition that negative errors are not taken into account. But as positive and negative errors are equally likely, the mean error in respect of negative errors is the same quantity, whence the mean error in respect of errors of both kinds is $.797884\sqrt{(\mu\div h)}$. This is usually called the *average* error. The mean error differs from the probable error in this respect, that it depends on the magnitude of individual errors, as well as on the proportion in which errors of different magnitudes occur. The probable error is independent of the magnitude.

148. When the quantity μ (the mean square of the errors) has been found from a series of observations, the precision, weight, probable error, and mean error, of a com-

ing observation of the same kind are found by supposing $h=1$ in the above expressions, and are respectively

precision	.	.	$=\sqrt{1 \div 2\mu}$
weight	.	.	$=1 \div 2\mu$
probable error	.	.	$=.674489\sqrt{\mu} \text{ []}$
mean error	.	.	$=.398942\sqrt{\mu}$

149. The preceding formulæ give the limits of the error to be feared in determining the value of a quantity from a series of observations, when the thing to be determined is that on which the observations are immediately made. We have now to apply the formulæ to the cases in which the quantity sought is not observed itself, but is a function of several others, which are separately determined by observation. The following problem is important :

Let u be a given function of a number of unknown quantities, $x, x', x'', \&c.$; it is required to assign the limits of the probable error in the determination of U , and the weight of the result, when values of x, x', x'' , found from observations independent of each other, and respectively affected with the probable errors $\varrho\sqrt{\mu}, \varrho\sqrt{\mu'}, \varrho\sqrt{\mu''}, \&c.$ ($\varrho = .674489$) are adopted instead of the true but unknown values of those quantities.

Let $u=f(x, x', x'', \&c.)$ be the given function, $\lambda, \lambda', \lambda'', \&c.$ observed values of $x, x', x'', \&c.$ and make $\lambda-x=e, \lambda'-x'=e', \lambda''-x''=e'', \&c.$ so that $e, e', e'', \&c.$ are the errors of observation, supposed to be so small that their squares may be rejected. Make $\frac{du}{dx}=a, \frac{du}{dx'}=a', \frac{du}{dx''}=a'',$

$\&c.$, then $a, a', a'',$ are given quantities; and on substituting $x+e, x'+e', x''+e''$ for $x, x', x'',$ respectively, in the equation $u=f(x, x', x'', \&c.)$, and supposing u to become

$u + E$ when the substitutions are made, so that E is the corresponding error of u , we have, on expanding u by Taylor's theorem,

$$E = ae + a'e' + a'e'' + \&c.$$

in respect of a single observation of each of the quantities. Taking the square of both sides of the equation, we have

$$E^2 = a^2e^2 + a'^2e'^2 + a''^2e''^2 + \&c. + 2aa'ee' + 2aa''ee'' + 2a'a'e'e'' + \&c.$$

Now since positive and negative errors are supposed equally probable, the sums of the products ee' , ee'' , $e'e''$, &c. or their mean values, become each $= 0$; therefore

$$\Sigma E^2 = a^2 \Sigma e^2 + a'^2 \Sigma e'^2 + a''^2 \Sigma e''^2 + \&c.$$

Taking the mean value of each of these sums, and observing that μ the mean value of Σe^2 is independent of the number of observations (142), and assuming M to be the mean value of ΣE^2 , we get

$$M = a^2\mu + a'^2\mu' + a''^2\mu'' + \&c.$$

This equation contains the solution of the problem, for all the functions of the error are given in terms of M . The probable error is $\cdot 674489\sqrt{M}$.

150. Let W be the weight of the determination, and $w, w', w'', \&c.$ the weights corresponding to $\mu, \mu', \mu'', \&c.$ then by the definition of weight, w is reciprocally proportioned to μ , and W to M ; and we have by substitution,

$$W = 1 \div \left(\frac{a^2}{w} + \frac{a'^2}{w'} + \frac{a''^2}{w''} + \&c. \right)$$

If the weights are supposed all equal, this becomes

$$W = \frac{w}{a^2 + a'^2 + a''^2 + \&c.}$$

Suppose the errors $e, e', e'', \&c.$ to be respectively multiplied by numbers proportional to the square roots of the weights, (which is equivalent to supposing all the observa-

tions to have the same degree of precision measured by $\sqrt{(\mu w)}$, then the value of M becomes

$$M = a^2 \mu w + a'^2 \mu' w' + a''^2 \mu'' w'' + \&c.$$

But w being reciprocally as μ , we have $\mu w = \mu' w' = \mu'' w''$, &c.
 $= 1$, therefore

$$W = \frac{1}{a^2 + a'^2 + a''^2 + \&c.}$$

SECTION X.

OF THE METHOD OF LEAST SQUARES.

151. In the determination of astronomical and physical elements from the data of observation, the thing which is actually observed is for the most part not the element which is sought to be determined, but a known function of that element. Thus, if V be a given function of X determined by the equation $V=F(X)$, the quantity observed may be a value of V , whilst the element sought to be determined is X . If the observation could give the value of V with absolute accuracy, then X would also be absolutely known; but as all observations are affected with certain errors of greater or less amount, owing to the imperfections of instruments or of sense, or the ever varying circumstances under which they are made, an exact value of X cannot be found from any single observation; and in order to obtain the utmost precision, it is necessary to employ a great number of observations, repeated under every variety of circumstance by which the result can be supposed to be affected.

152. The observed quantity V , instead of being a function of a single element X , may be a function of several elements X , Y , Z , &c.; for example, V may be the position of a planet, in which case it is a function of the six elements of the orbit, for the determination of which the

observation is made. Each observation gives rise to an equation of this form, $V=F(X, Y, Z, \&c.)$; therefore when the number of equations is just equal to the number of unknown quantities, the problem is determinate; and supposing F to be an algebraic function, the values of $X, Y, Z, \&c.$ may be found by the ordinary methods of elimination. If the number of equations is less than the number of unknown quantities, the problem is indeterminate; but if greater, it may be said to be more than determinate, inasmuch as the equations may be combined in an infinite number of ways, each distinct combination giving a different value of the elements. It therefore becomes a question of the utmost importance to the perfection of the sciences of observation, to assign the particular combination which gives the most advantageous results, or values of $X, Y, Z, \&c.$ affected with the smallest probable errors.

153. As approximate values of the elements are in all cases either already known, or can be easily found, the object of accumulating observations is the correction of the approximate values. Let V be the true value of the thing observed, V_o an approximate value, however found, X the true value of the element sought, X_o an approximate value, corresponding to V_o , so that we have the two equations $V=F(X), V_o=F(X_o)$; also, let the observed value of V in any observation be L , and make

$$v=V-L, \quad l=V_o-L,$$

then v is the true but unknown error of the observation, and l its reputed error, that is to say, the difference between the computed value of the function and the result of the observation. Now if we assume x to represent the true correction of the approximate element, so that $X=X_o+x$, then, on substituting X_o+x for X in the function F , we

get $V=F(X_0+x)$; whence, expanding the function by Taylor's theorem, and rejecting terms multiplied by x^2 and higher powers of x , because x is a very small quantity

$$V=V_0+\frac{dV_0}{dX}x.$$

Let us now denote the differential coefficient, which is a known quantity, by a ; then, observing that $V-V_0=v-l$, the equation becomes $v=l+ax$; that is to say, the true error of the observation is a linear function of the correction of the element.

154. In like manner, when there are several elements, X, Y, Z , &c.; on making $\frac{dV_0}{dX}=a, \frac{dV_0}{dY}=b, \frac{dV_0}{dZ}=c$, &c. a single observation furnishes the equation

$$v=l+ax+by+cz+\&c.,$$

and a series of observations, whose errors are respectively v, v', v'' , &c. gives a system of linear equations equal in number to the number of observations; namely,

$$\begin{aligned} v &= l + ax + by + cz + \&c. \\ v' &= l' + a'x + b'y + c'z + \&c. \\ v'' &= l'' + a''x + b''y + c''z + \&c. \\ &\&c. \end{aligned} \tag{1}$$

and the object is to give such values to x, y, z , &c. that the errors v, v', v'' , &c. in respect of the whole of the observations, shall be the least possible. The equations being supposed independent of each other, if their number is just equal to that of the unknown quantities, the errors v, v', v'' , &c. can be made all zero; but if, as is usually the case, there are more equations than unknown quantities, it is impossible by any means whatever to annihilate the whole of them, and therefore all that can be accomplished is to find the system of values of x, y, z , &c. which most nearly,

and with the greatest probability, satisfies the whole of the equations. If the observations are not all equally good, the equations are supposed to be each multiplied by a number proportional to the square root of the presumed weight of the observation on which it depends, in order that they may all have the same degree of precision.

155. As the question is to find the most probable values of x, y, z , &c. the first thing necessary is to express each of these elements in terms of the observations. Suppose k, k', k'' , &c. to be a system of indeterminate quantities, independent of x, y, z , &c. and let the first of the above conditional equations be multiplied by k , the second by k' , the third by k'' , and so on; then adding the products, if k, k', k'' , &c. be determined so as to make the coefficient of x equal to unit, and those of y, z , &c. each equal to 0; that is to say, so as to satisfy the equations

$$\begin{aligned} ka + k'a' + k''a'' + \&c. &= 1 \\ kb + k'b' + k''b'' + \&c. &= 0 \\ kc + k'c' + k''c'' + \&c. &= 0 \\ \&c. & \end{aligned} \quad (2)$$

we shall then have $x = K + kv + k'v' + k''v'' + \&c.$ where K is a quantity independent of v, v', v'' , &c. Hence x is found $= K$, with an error $= kv + k'v' + k''v'' + \&c.$; and the *weight* of the determination, by the formula in (150), is

$\frac{1}{k^2 + k'^2 + k''^2 + \&c.}$. The weight of the determination is

consequently greater in proportion as $k^2 + k'^2 + k''^2 + \&c.$ is smaller; and hence of all the possible systems of indeterminate coefficients, k, k', k'' , &c. which satisfy the equations (1), the system which gives the most probable value of x , or the most advantageous result, is that for which $k^2 + k'^2 + k''^2 + \&c.$ is an absolute minimum.

156. We have now to find, in terms of known quantities, values of the indeterminate coefficients $h, h', h'',$ &c. which satisfy the condition of the minimum. For the sake of abridging, let us denote the aggregate of the products $aa + a'a' + a''a'' + \&c.$ by $S(aa)$, that of $ab + a'b' + a''b'' + \&c.$ by $S(ab)$, and so on, and also assume

$$\begin{aligned}\xi &= av + a'v' + a''v'' + \&c. \\ \eta &= bv + b'v' + b''v'' + \&c. \\ \zeta &= cv + c'v' + c''v'' + \&c.\end{aligned}\tag{3}$$

On substituting in these equations the values of $v, v', v'',$ &c. given by the equations (1), there results

$$\begin{aligned}\xi &= S(al) + xS(aa) + yS(ab) + zS(ac) + \&c. \\ \eta &= S(bl) + xS(ab) + yS(bb) + zS(bc) + \&c. \\ \zeta &= S(cl) + xS(ac) + yS(bc) + zS(cc) + \&c.\end{aligned}\tag{4}$$

a system of equations equal in number to the number of elements $x, y, z,$ &c. and from which, consequently, those elements would be determined absolutely if the observations were perfectly exact, that is, if the errors $v, v', v'',$ &c. were individually zero, and consequently $\xi, \eta, \zeta,$ &c., were each zero. On eliminating $y, z,$ &c. from the last system, the value of x is given in terms of $\xi, \eta, \zeta,$ and known quantities by a linear equation of the following form :

$$x = A + f\xi + g\eta + h\zeta + \&c.\tag{5}$$

where $f, g, h,$ &c. are co-efficients independent of $x, y, z,$ &c. and also of $\xi, \eta, \zeta,$ &c.

If we now substitute in equation (5) the values $\xi, \eta, \zeta,$ &c. given by equations (3), and also assume

$$\begin{aligned}a &= fa + gb + hc + \&c. \\ a' &= f'a' + g'b' + h'c' + \&c. \\ a'' &= f''a'' + g''b'' + h''c'' + \&c.,\end{aligned}\tag{6}$$

we shall have by addition

$$x = A + av + a'v' + a''v'' + \&c. ;$$

whence it appears that $a, a', a'',$ &c. are a system of multi-

pliers by which $y, z, \&c.$ are eliminated from equations (1); they must therefore satisfy the equations (2), whence

$$\begin{aligned}aa + a'a' + a''a'' + \&c. &= 1 \\ab + a'b' + a''b'' + \&c. &= 0 \\ac + a'c' + a''c'' + \&c. &= 0.\end{aligned}\quad (7)$$

Subtracting these from the equations (2) we obtain

$$\begin{aligned}0 &= (k-a)a + (k'-a')a' + (k''-a'')a'' + \&c. \\0 &= (k-a)b + (k'-a')b' + (k''-a'')b'' + \&c. \\0 &= (k-a)c + (k'-a')c' + (k''-a'')c'' + \&c.,\end{aligned}$$

on multiplying which respectively by f, g, h , and adding the products, we get by reason of the equations (6),

$$0 = (k-a)a + (k'-a')a' + (k''-a'')a'' + \&c.$$

This equation may be put under the form $k^2 + k'^2 + k''^2 + \&c.$

$$= a^2 + a'^2 + a''^2 + \&c. + (k-a)^2 + (k'-a')^2 + (k''-a'')^2 + \&c.,$$

from which it is evident that $k^2 + k'^2 + k''^2 + \&c.$ will be a minimum when $k=a, k'=a', k''=a'', \&c.$ Hence it follows that the most probable value of x which can be deduced from the equations (1), is $x=A$; and by (150) the weight of the determination is $1 \div (aa + a'a' + a''a'' + \&c.) = 1 \div S(aa).$

This quantity $S(aa)$ is equal to f the co-efficient of ξ in the equation (5); for on multiplying the first of equations (7), by f , the second by g , and the third by h , and adding the products, we obtain by reason of equations (6),

$$aa + a'a' + a''a'' + \&c. = f.$$

157. The method explained in the two last paragraphs of determining the most advantageous combination of a system of linear equations, of the form of those in (154), is given by Gauss in his *Theoria Combinationis Observationum erroribus minimis obnoxia*, (Gottingen, 1823). The practical rule to which it leads is as follows: Having given a near value V of a function of several elements, $X, Y, Z,$

&c. and also a series $L, L', L'',$ &c. of observed values of V , make $(V-L)\sqrt{w}=v, (V-L')\sqrt{w'}=v', (V-L'')\sqrt{w'}=v'',$ &c. and form the equations in (1). From these equations (4) are easily deduced; and from these, again, by elimination, are found the values of $x, y, z,$ &c. the corrections of the approximate elements $X, Y, Z,$ &c., in equations of the form (5), which, for the sake of symmetry, may be thus written :

$$x=A+(aa)\xi+(a\beta)\eta+(a\gamma)\zeta+\&c.$$

$$y=B+(a\beta)\xi+(\beta\beta)\eta+(\beta\gamma)\zeta+\&c.$$

$$z=C+(a\gamma)\xi+(\beta\gamma)\eta+(\gamma\gamma)\zeta+\&c.$$

then the most probable values of $x, y, z,$ &c. are respectively $A, B, C,$ &c.; the weights of the determinations respectively

$\frac{1}{(aa)}, \frac{1}{(\beta\beta)}, \frac{1}{(\gamma\gamma)},$ &c. ; and the probable errors of the several determinations are $\rho\sqrt{(aa)}, \rho\sqrt{(\beta\beta)}, \rho\sqrt{(\gamma\gamma)},$ &c., where $\rho=.476936$.

158. The values of $x, y, z,$ &c. now deduced are obtained immediately, by supposing the sum of the squares of the errors of observation to be a minimum. Thus, forming the squares of the equations (1), and making $\Omega=v^2+v'^2+v''^2+\&c.,$ the differentiation of Ω in respect of each of the variables $x, y, z,$ &c. produces the quantities denoted in (156), by $\xi, \eta, \zeta,$ &c. that is to say, it gives

$$\frac{d\Omega}{dx}=2\xi, \frac{d\Omega}{dy}=2\eta, \frac{d\Omega}{dz}=2\zeta, \&c.$$

therefore if Ω be a minimum, $\xi, \eta, \zeta,$ become severally zero, and the equations (4) give by elimination, $x=A, y=B, z=C,$ where $A, B,$ and C denote the same quantities as above. Now from equation (5) the general value of x is

$$x=A+(aa)\xi+(a\beta)\eta+(a\gamma)\zeta,$$

and the most probable value being $x=A,$ it follows that

the most probable values of the corrections x, y, z , are found by making the differential coefficients of Ω equal to zero, that is, by making $v^2 + v'^2 + v''^2 + \&c.$ an absolute minimum. Hence this method of combining equations of condition is called the *method of least squares*; and it follows from the preceding analysis, that it gives the most probable values of the corrections, or the most advantageous results.

159. As an example, let us suppose there is only one unknown element X , of which X_o is known to be an approximate value, and $L, L', L'', \&c.$ are observed values, the weights of which are respectively proportional to $w, w', w'', \&c.$ and that it is required to determine the most probable value of X from the observations, and also the weight of the determination. Make $(X - L)\sqrt{w} = v, (X_o - L)\sqrt{w} = l$, and let x be the correction of X , so that $X = X_o - x$. On substituting this in $(X - L)\sqrt{w} = v$, we have $(X_o - x - L)\sqrt{w} = v$, or $v = l\sqrt{w} - x\sqrt{w}$. Each observation gives a similar equation, and the equations (1) in (154) consequently become

$$\begin{aligned} v &= l\sqrt{w} - x\sqrt{w} \\ v' &= l'\sqrt{w'} - x\sqrt{w'} \\ v'' &= l''\sqrt{w''} - x\sqrt{w''}, \&c. \end{aligned}$$

therefore, multiplying each by the coefficient of its own x , we have $\xi = S(lw) - xS(w)$, whence consequently $x = S(lw) \div S(w) - \xi \div S(w)$, that is to say, the most probable value of x is

$$x = \frac{lw + l'w' + l''w'' + \&c.}{w + w' + w'' + \&c.},$$

and the weight of the determination is proportional to the reciprocal of $w + w' + w'' + \&c.$

Since $X - L = x - l$, we have also

$$X = \frac{Lw + L'w' + L''w'' + \&c.}{w + w' + w'' + \&c.},$$

whence this proposition: If a series of values of an element

are found from observations which have not all the same degree of precision, the most probable value of the element is found by multiplying each observation by a number proportional to its weight, and dividing the sum of the products by the sum of the weights; and the comparative weight of the result is unit divided by the sum of all the weights.

If the weights be all equal, and the number of the observations be h , then $X = (L + L' + L'' + \&c.) \div h$; that is to say, the average of a series of equally good observations gives the most probable value. The average may, therefore be considered as a particular case of the method of least squares.

160. To illustrate the method of proceeding when there are several elements to be corrected from the observations, we shall take the following numerical example from Gauss (*Theoria Motus*). Suppose there are three elements, and that three observations, of equal weight, have given the equations $x - y + 2z = 3$, $3x + 2y - 5z = 5$, $4x + y + 4z = 21$; and that a fourth observation, of which the relative weight is one-fourth, or its precision one-half, of that of the others has given $-2x + 6y + 6z = 28$. The first step is to reduce this last equation to the same standard of weight with the others, for which purpose it must be multiplied by $\frac{1}{2}$; it then becomes $-x + 3y + 3z = 14$. Now, as x , y , and z cannot be determined so as to satisfy four independent equations, we suppose each observation, or equation, to be affected with an error v , and accordingly obtain the following system of equations, corresponding to equations (1), viz.:

$$\begin{aligned} v &= -3 + x - y + 2z \\ v' &= -5 + 3x + 2y - 5z \\ v'' &= -21 + 4x + y + 4z \\ v''' &= -14 - x + 3y + 3z, \end{aligned}$$

from which the most probable values of x , y , and z are to be deduced. Let each equation be multiplied by the coefficient of its own x , taken with its proper sign, namely, the first by 1, the second by 3, the third by 4, and the fourth by -1 ; the results added together give the value of ξ , namely, $\xi = -88 + 27x + 6y$. In like manner, let the first be multiplied by -1 , the second by 2, the third by 1, and the fourth by 3, the sum of the products will give η . Lastly, let the equations be multiplied respectively by the coefficients of z , and the sum of the products made equal to ζ ; we have then the following equations corresponding to the equations (4)

$$\xi = -88 + 27x + 6y + 0$$

$$\eta = -70 + 6x + 15y + z$$

$$\zeta = -107 + 0 + y + 54z.$$

From these we get by elimination

$$19899x = 49154 + 809\xi - 324\eta + 6\zeta$$

$$737y = 2617 - 12\xi + 54\eta - \zeta$$

$$39798z = 76242 + 12\xi - 54\eta + 1473\zeta,$$

whence (157) A , B , C , the most probable values of x , y , z , are respectively

$$A = \frac{49154}{19899} = 2.470, B = \frac{2617}{737} = 3.551, C = \frac{76242}{39798} = 1.916;$$

and the relative weights w , w' , w'' , are respectively

$$w = \frac{19899}{809} = 24.6, w' = \frac{737}{54} = 13.6, w'' = \frac{39798}{1473} = 27.0,$$

whence the probable errors ($.476936 \div \sqrt{w}$) are respectively
0.96, .129, .092.

The method of least squares, to which modern astronomy is indebted for much of its precision, was first proposed by Legendre, in his *Nouvelles Méthodes pour la Détermination les Orbites des Comètes*, (Paris, 1806,) merely as

means of avoiding inconvenience and uncertainty arising from the want of a uniform and determinate method of combining numerous equations of condition, and without reference to the theory of probability. The same method, however, had previously been discovered by Gauss, and a demonstration of it, deduced from the general theory of chances, was given by him in his *Theoria Motus*, (1809.) It may be shewn in various ways, that this method of combination gives values of the unknown quantities affected with the smallest probable errors; but it is to be observed, that all the demonstrations are subordinate to the hypothesis, that positive and negative errors of equal magnitude are equally probable, or that the average of a large number of results gives the most probable value, and consequently that the function which represents the probability of an error has the form assigned to it in (132).

The limits of this article will not permit us to enter into further details respecting the applications of the method of least squares. On the general theory of the probable errors of results deduced from observation, and the most advantageous methods of combining equations of condition, the reader may consult the *Théorie Analytique des Probabilités* of Laplace; the *Theoria Motus* of Gauss; the *Theoria Combinationis Observationum*, and the *Supplementum Theoriæ Combinationis*, &c. (Göttingen, 1828) of the same author; the *Recherches sur la Probabilité des Jugemens*, with the two Memoirs of Poisson in the *Connaissance des Temps* for 1827 and 1832; and three masterly papers, by Mr. Ivory, in the *Philosophical Magazine* for 1825. In the volumes of the *Berliner Astronomisches Jahrbuch* for 1833, 1834, and 1835, M. Encke has treated the subject at great length, and given a number of formulæ calculated

to facilitate the labours of the computer. We may also refer, in conclusion, to a very recent and remarkable disquisition on the theory of probable errors, by the celebrated astronomer Bessel, forming Nos. 358 and 359 of Schumacher's *Astronomische Abhandlungen*, Altona, October 1838.

TABLE

OF THE

VALUES OF THE INTEGRAL

$$\Theta = \frac{2}{\sqrt{\pi}} \int_0^{\tau} e^{-t^2} dt = 1 - \frac{2}{\sqrt{\pi}} \int_{\tau}^{\infty} e^{-t^2} dt,$$

for intervals each = 0.1, from $\tau=0$ to $\tau=3$, with their first and second differences.

τ	Θ	Δ	Δ^2	τ	Θ	Δ	Δ^2
0.00	0.0000000	112833	22	0.30	.3286267	102814	636
0.01	.0112833	112811	45	0.31	.3389081	102178	652
0.02	.0225644	112766	67	0.32	.3491259	101526	667
0.03	.0338410	112699	90	0.33	.3592785	100859	684
0.04	.0451109	112609	112	0.34	.3693644	100175	698
0.05	.0563718	112497	135	0.35	.3793819	99477	714
0.06	.0676215	112362	158	0.36	.3893296	98763	729
0.07	.0788577	112204	179	0.37	.3992059	98034	742
0.08	.0900781	112025	201	0.38	.4090093	97292	755
0.09	.1012806	111824	224	0.39	.4187385	96537	769
0.10	.1124630	111600	246	0.40	.4283922	95768	782
0.11	.1236230	111354	267	0.41	.4379690	94986	795
0.12	.1347584	111087	288	0.42	.4474676	94191	807
0.13	.1458671	110799	310	0.43	.4568867	93384	817
0.14	.1569470	110489	331	0.44	.4662251	92567	830
0.15	.1679959	110158	352	0.45	.4754818	91737	840
0.16	.1790117	109806	372	0.46	.4846555	90897	851
0.17	.1899923	109434	393	0.47	.4937452	90046	861
0.18	.2009357	109041	414	0.48	.5027498	89185	869
0.19	.2118398	108627	434	0.49	.5116683	88316	878
0.20	.2227025	108193	453	0.50	.5204999	87438	888
0.21	.2335218	107740	473	0.51	.5292437	86550	896
0.22	.2442958	107267	492	0.52	.5378987	85654	903
0.23	.2550225	106775	512	0.53	.5464641	84751	910
0.24	.2657000	106263	529	0.54	.5549392	83841	917
0.25	.2763263	105734	549	0.55	.5633233	82924	923
0.26	.2868997	105185	567	0.56	.5716157	82001	930
0.27	.2974182	104618	584	0.57	.5798158	81071	935
0.28	.3078800	104034	601	0.58	.5879229	80136	940
0.29	.3182834	103433	619	0.59	.5959365	79196	945

τ	Θ	Δ	Δ^2	τ	Θ	Δ	Δ^2
0.60	0.6038561	78251	949	1.00	.8427008	41097	822
0.61	.6116812	77302	953	1.01	.8468105	40275	813
0.62	.6194114	76349	955	1.02	.8508380	39462	805
0.63	.6270463	75394	959	1.03	.8547842	38657	796
0.64	.6345857	74435	962	1.04	.8586499	37861	786
0.65	.6420292	73473	963	1.05	.8624360	37075	777
0.66	.6493765	72510	965	1.06	.8661435	36297	768
0.67	.6566275	71545	966	1.07	.8697732	35529	760
0.68	.6637820	70579	968	1.08	.8733261	34769	749
0.69	.6708399	69611	967	1.09	.8768030	34020	740
0.70	.6778010	68644	968	1.10	.8802050	33280	731
0.71	.6846654	67676	968	1.11	.8835330	32549	721
0.72	.6914330	66708	966	1.12	.8867879	31828	712
0.73	.6981038	65742	966	1.13	.8899707	31116	701
0.74	.7046780	64776	965	1.14	.8930823	30415	691
0.75	.7111556	63811	962	1.15	.8961238	29724	682
0.76	.7175367	62849	961	1.16	.8990962	29042	672
0.77	.7238216	61888	957	1.17	.9020004	28370	661
0.78	.7300104	60931	956	1.18	.9048374	27709	652
0.79	.7361035	59975	952	1.19	.9076083	27057	642
0.80	.7421010	59023	948	1.20	.9103140	26415	631
0.81	.7480033	58075	945	1.21	.9129555	25784	622
0.82	.7538108	57130	941	1.22	.9155339	25162	611
0.83	.7595238	56189	936	1.23	.9180501	24551	602
0.84	.7651427	55253	931	1.24	.9205052	23949	591
0.85	.7706680	54322	926	1.25	.9229001	23358	581
0.86	.7761002	53396	921	1.26	.9252359	22777	571
0.87	.7814398	52475	916	1.27	.9275136	22206	561
0.88	.7866873	51559	909	1.28	.9297342	21645	552
0.89	.7918432	50650	904	1.29	.9318987	21093	541
0.90	.7969082	49746	897	1.30	.9340080	20552	532
0.91	.8018828	48849	891	1.31	.9360632	20020	522
0.92	.8067677	47958	883	1.32	.9380652	19498	511
0.93	.8115635	47075	877	1.33	.9400150	18987	502
0.94	.8162710	46198	870	1.34	.9419137	18485	493
0.95	.8208908	45328	861	1.35	.9437622	17992	482
0.96	.8254236	44467	855	1.36	.9455614	17510	474
0.97	.8298703	43612	846	1.37	.9473124	17036	463
0.98	.8342315	42766	839	1.38	.9490160	16573	455
0.99	.8385081	41927	830	1.39	.9506733	16118	445

τ	Θ	Δ	Δ^2	τ	Θ	Δ	Δ^2
1.40	0.9522851	15673	435	1.80	.9890905	4340	154
1.41	.9538524	15238	427	1.81	.9895245	4186	150
1.42	.9553762	14811	418	1.82	.9899431	4036	144
1.43	.9568573	14393	409	1.83	.9903467	3892	141
1.44	.9582966	13984	399	1.84	.9907359	3751	136
1.45	.9596950	13585	391	1.85	.9911110	3615	133
1.46	.9610535	13194	382	1.86	.9914725	3482	127
1.47	.9623729	12812	374	1.87	.9918207	3355	124
1.48	.9636541	12438	365	1.88	.9921562	3231	120
1.49	.9648979	12073	357	1.89	.9924793	3111	116
1.50	.9661052	11716	349	1.90	.9927904	2995	112
1.51	.9672768	11367	340	1.91	.9930899	2883	108
1.52	.9684135	11027	332	1.92	.9933782	2775	106
1.53	.9695162	10695	325	1.93	.9936557	2669	101
1.54	.9705857	10370	316	1.94	.9939226	2568	99
1.55	.9716227	10054	309	1.95	.9941794	2469	95
1.56	.9726281	9745	301	1.96	.9944263	2374	91
1.57	.9736026	9444	294	1.97	.9946637	2283	89
1.58	.9745470	9150	286	1.98	.9948920	2194	85
1.59	.9754620	8864	279	1.99	.9951114	2109	84
1.60	.9763484	8585	273	2.00	.9953223	2025	78
1.61	.9772069	8312	264	2.01	.9955248	1947	79
1.62	.9780381	8048	259	2.02	.9957195	1868	73
1.63	.9788429	7789	251	2.03	.9959063	1795	72
1.64	.9796218	7538	245	2.04	.9960858	1723	69
1.65	.9803756	7293	238	2.05	.9962581	1654	67
1.66	.9811049	7055	231	2.06	.9964235	1587	65
1.67	.9818104	6824	226	2.07	.9965822	1522	61
1.68	.9824928	6598	220	2.08	.9967344	1461	61
1.69	.9831526	6378	212	2.09	.9968805	1400	57
1.70	.9837904	6166	208	2.10	.9970205	1343	55
1.71	.9844070	5958	201	2.11	.9971548	1288	54
1.72	.9850028	5757	196	2.12	.9972836	1234	51
1.73	.9855785	5561	190	2.13	.9974070	1183	50
1.74	.9861346	5371	185	2.14	.9975253	1133	47
1.75	.9866717	5186	179	2.15	.9976386	1086	47
1.76	.9871903	5007	175	2.16	.9977472	1039	45
1.77	.9876910	4832	168	2.17	.9978511	994	40
1.78	.9881742	4664	165	2.18	.9979505	954	41
1.79	.9886406	4499	159	2.19	.9980459	913	41

τ	Θ	Δ	Δ^2	τ	Θ	Δ	Δ^2
2.20	0.9981372	872	37	2.60	.9997640	127	6
2.21	.9982244	835	36	2.61	.9997767	121	6
2.22	.9983079	799	35	2.62	.9997888	115	6
2.23	.9983878	764	33	2.63	.9998003	109	6
2.24	.9984642	731	33	2.64	.9998112	103	5
2.25	.9985373	698	30	2.65	.9998215	98	5
2.26	.9986071	668	30	2.66	.9998313	93	5
2.27	.9986739	638	29	2.67	.9998406	88	4
2.28	.9987377	609	27	2.68	.9998494	84	5
2.29	.9987986	582	26	2.69	.9998578	79	4
2.30	.9988568	556	25	2.70	.9998657	75	4
2.31	.9989124	531	24	2.71	.9998732	71	4
2.32	.9989655	507	23	2.72	.9998803	67	4
2.33	.9990162	484	23	2.73	.9998870	63	2
2.34	.9990646	461	20	2.74	.9998933	61	4
2.35	.9991107	441	21	2.75	.9998994	57	3
2.36	.9991548	420	19	2.76	.9999051	54	3
2.37	.9991968	401	19	2.77	.9999105	51	3
2.38	.9992369	382	18	2.78	.9999156	48	2
2.39	.9992751	364	17	2.79	.9999204	46	3
2.40	.9993115	347	16	2.80	.9999250	43	2
2.41	.9993462	331	16	2.81	.9999293	41	3
2.42	.9993793	315	15	2.82	.9999334	38	1
2.43	.9994108	300	14	2.83	.9999372	37	3
2.44	.9994408	286	14	2.84	.9999409	34	1
2.45	.9994694	272	12	2.85	.9999443	33	2
2.46	.9994966	260	14	2.86	.9999476	31	2
2.47	.9995226	246	11	2.87	.9999507	29	2
2.48	.9995472	235	12	2.88	.9999536	27	1
2.49	.9995707	223	10	2.89	.9999563	26	2
2.50	.9995930	213	11	2.90	.9999589	24	1
2.51	.9996143	202	10	2.91	.9999613	23	1
2.52	.9996345	192	9	2.92	.9999636	22	1
2.53	.9996537	183	10	2.93	.9999658	21	2
2.54	.9996720	173	8	2.94	.9999679	19	1
2.55	.9996893	165	8	2.95	.9999698	18	1
2.56	.9997058	157	8	2.96	.9999716	17	...
2.57	.9997215	149	8	2.97	.9999733
2.58	.9997364	141	6	
2.59	.9997505	135	8	3.00	.9999779

ERRATA.

Page 40, note, for *Essai sur les Probabilités*, read *Théorie Analytique des Probabilités*.

— 43, for *adopt*, read *adapt*.

— 49, line 13, read beginning *with*. In the note, read *Traité Élémentaire*.

— 89, line 9, for $m-m$, read $m+m'$.

121-

7 DAY USE

RETURN TO DESK FROM WHICH BORROWED
ASTRONOMY, MATHEMATICS-
STATISTICS LIBRARY

This publication is due on the LAST DATE
and HOUR stamped below.

MAR 17 1993

4/25

NON-CIRCULATING

RB 17-50m-7,'65
(F5759s10)4188

General Library
University of California
Berkeley

Page 22

Page 37. art 19. Br.

Page 47—

Page 54. art 28

0
Art 20. Page 60

U.C. BERKELEY LIBRARIES



C036286147

609735

Q4

273

G34

Math
dept.

UNIVERSITY OF CALIFORNIA LIBRARY

